

RESULT. 2

ARTICULANT: CERNETIL, DOUGLAS F.  
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE

```

1  TITLE OF INVENTION:  RECEPTOR HEK
2  NUMBER OF SEQUENCES:  4
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  IMMUNEX CORPORATION
5  STREET:  51 UNIVERSITY STREET
6  CITY:  SEATTLE
7  STATE:  WASHINGTON
8  COUNTRY:  USA
9  ZIP:  98101
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE:  Floppy disk
12 COMPUTER:  Apple Macintosh
13 OPERATING SYSTEM:  Apple System 7.1
14 SOFTWARE:  Microsoft Word for Apple, Version 5.1a
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER:  US/08/240,124
17 FILING DATE:
18 CLASSIFICATION:  435
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER:  US 08/161,132
21 FILING DATE:  03-DEC-1993
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER:  US 08/114,426
24 FILING DATE:  30-AUG-1993
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER:  US 08/109,745
27 FILING DATE:  20-AUG-1993
28 ATTORNEY/AGENT INFORMATION:
29 NAME:  SESE, KATHRYN A.
30 REGISTRATION NUMBER:  32,172
31 REFERENCE/DOCKET NUMBER:  2814-C
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE:  (206) 587-0430
34 TELEFAX:  (206) 233-0644
35 TELEX:  756822
36 INFORMATION FOR SEQ ID NO:  1:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH:  1037 base pairs
39 TYPE:  nucleic acid
40 STRANDEDNESS:  single
41 TOPOLOGY:  linear
42 MOLECULE TYPE:  cDNA to mRNA
43 HYPOTHETICAL:  NO
44 ANTI-SENSE:  NO
45 IMMEDIATE SOURCE:
46 CLONE:  hek-L A2
47 FEATURE:
48 NAME/KEY:  CDS
49 LOCATION:  83..799
50 FEATURE:
51 NAME/KEY:  sig_peptide
52 LOCATION:  83..139
53 FEATURE:
54 NAME/KEY:  mat_peptide
55 LOCATION:  140..796
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925

```

Query Match	52.7%	Score 917.8;	DB 1;	Length 1037;
Best Local Similarity	99.8%	Pred. No. 2.4e-179;		
Matches 919; Conservative	0;	Mismatches 2;	Indels 0;	

[illegible]

```

COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,943
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/240,124
FILING DATE: 09-MAY-1994
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 387-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hex-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-08-453-943-1

```

Query Match	52.7%	Score 917.8;	DB 1;	Length 1037;
Best Local Similarity	99.8%;	Pred. NO. 2.4e-179;		
Matches 919:	Conservative	0: Mismatches	2: Indels	0: Gaps

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 10:44:50 ; Search time 65.87 Seconds  
(without alignments)  
625.061 Million cell updates/sec

Title: US-09-733-756-2

Perfect score: 1301

Sequence: 1 MAAPLPLLLLVVPLPL.....REHPLAVGIAFTLMTFLAS 238

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP\_archaea:\*
- 2: SP\_bacteria:\*
- 3: SP\_fungi:\*
- 4: SP\_human:\*
- 5: SP\_invertebrate:\*
- 6: SP\_mammal:\*
- 7: SP\_mhc:\*
- 8: SP\_organelle:\*
- 9: SP\_phase:\*
- 10: SP\_plant:\*
- 11: SP\_rodent:\*
- 12: SP\_virus:\*
- 13: SP\_vertebrate:\*
- 14: SP\_unclassified:\*
- 15: SP\_rvirus:\*
- 16: SP\_bacteriap:\*
- 17: SP\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	727	55.9	219	13 Q90YC5	Q90YC5 brachydanio
2	586	45.0	118	11 Q923G4	Q923G4 rattus norv
3	424.5	32.6	202	13 Q987Z1	Q987Z1 gallus gall
4	406	31.2	88	13 Q9PTD0	Q9PTD0 ctenophorus
5	406	31.2	206	11 Q9CZS8	Q9CZS8 mus musculus
6	373	28.7	229	13 Q934J1	Q934J1 brachydanio
7	367	28.2	205	11 Q9D7K8	Q9D7K8 mus musculus
8	338	26.0	102	11 Q9WUE7	Q9WUE7 rattus norv
9	268.5	20.6	93	13 Q9PTD1	Q9PTD1 ctenophorus
10	247.5	19.0	80	13 Q42304	Q42304 brachydanio
11	208	16.0	42	11 Q9CRL3	Q9CRL3 mus musculus
12	178	13.7	70	13 Q90ZG7	Q90ZG7 brachydanio
13	166.5	12.8	331	13 Q90Z31	Q90Z31 brachydanio
14	164.5	12.6	327	13 Q9PT69	Q9PT69 xenopus lae
15	162.5	12.5	333	13 Q9PUJ4	Q9PUJ4 gallus gall
16	145	11.1	652	5 Q9V4E1	Q9V4E1 drosophila

17	143.5	11.0	348	5	O44516	O44516 caenorhabdi
18	138.5	10.6	279	5	Q9U474	Q9U474 caenorhabdi
19	138	10.6	334	13	Q90Z32	Q90Z32 brachydanio
20	124.5	9.6	341	13	Q90Z33	Q90Z33 brachydanio
21	100	7.7	377	3	O74823	O74823 schizosacch
22	96	7.4	1199	5	Q9V9Y3	Q9V9Y3 drosophila
23	92	7.1	855	15	Q9YKQ0	Q9YKQ0 human immun
24	88.5	6.8	441	4	Q9NQ99	Q9NQ99 homo sapien
25	88.5	6.8	709	4	Q9NRN2	Q9NRN2 homo sapien
26	88.5	6.8	894	4	Q9UMZ9	Q9UMZ9 homo sapien
27	87.5	6.7	765	11	Q91ZV1	Q91ZV1 mus musculu
28	87.5	6.7	765	11	Q91V98	Q91V98 mus musculu
29	87.5	6.7	894	4	Q9BZH4	Q9BZH4 homo sapien
30	87	6.7	817	5	Q93560	Q93560 caenorhabdi
31	86	6.6	438	11	Q9R262	Q9R262 mus musculu
32	86	6.6	442	11	Q9Z2L7	Q9Z2L7 mus musculu
33	85.5	6.6	709	5	O16783	O16783 caenorhabdi
34	84	6.5	855	15	Q9YKQ4	Q9YKQ4 human immun
35	84	6.5	1027	5	Q17657	Q17657 caenorhabdi
36	83.5	6.4	457	3	Q05672	Q05672 saccharomyc
37	83.5	6.4	473	5	Q961A1	Q961A1 drosophila
38	83.5	6.4	488	3	O07619	O07619 saccharomyc
39	83.5	6.4	502	5	Q9W0V9	Q9W0V9 drosophila
40	83.5	6.4	561	5	Q9V9N4	Q9V9N4 drosophila
41	83.5	6.4	910	11	Q9JIL3	Q9JIL3 rattus norv
42	83	6.4	1342	5	Q9GPP6	Q9GPP6 drosophila
43	83	6.4	2206	12	Q9DIF1	Q9DIF1 human polio
44	82.5	6.3	304	4	Q9HBC3	Q9HBC3 homo sapien
45	82.5	6.3	463	4	Q13648	Q13648 homo sapien

ALIGNMENTS

RESULT	1
Q90YC5	
ID	Q90YC5 PRELIMINARY; PRT; 219 AA.
AC	Q90YC5;
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	EPHRIN-A3.
GN	EPHRIN-A3.
OS	Brachydanio rerio (Zebrafish) (Zebra danio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC	Cypriniformes; Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21412237; PubMed=11520665;
RA	Hirate Y., Mieda M., Harada T., Yamasu K., Okamoto H.;
RT	"Identification of ephrin-A3 and novel genes specific to the midbrain-
RT	MHB in embryonic zebrafish by ordered differential display.";
RL	Mech. Dev. 107:83-96(2001).
DR	EMBL; AB051678; BAB55891.1; .
SQ	SEQUENCE 219 AA; 25146 MW; 7191927E03F8EA01 CRC64;

Query Match	55.9%; Score 727; DB 13; Length 219;
Best Local Similarity	61.9%; Pred. No. 1.4e-67;
Matches 140; Conservative	25; Mismatches 43; Indels 18; Gaps 3;
Qy	1 MAAPLPLLLLVVPLPLLAQPGGALGNHAYWNSSNOHLRREGYTVQVNVNDYLD 60
Db	1 MALYALCLFLLTCTCNALVTAA-----RHAVHNSNLLLRKGYTLQVNVNDYLD 53
Qy	61 IYCPHNSVGVGPGAGGGAGAEQYVLYMVSRYRTCNASQGPKRWECNRPAPHSP1 120
Db	54 IYCPHNSV-----QRGIAEQYVLYMVSRYRTCDPOLGPKRWECNRPAPHAP1 104
Qy	121 KFSEKFORYSAFSLGVEHAGHEYYYISTPTNLHLKCLRMKVFVCCASTSGSEKPVPT 180

DR PROSITE; PS01299; EPHRN; 1.  
FT NON\_TER 1  
SQ SEQUENCE 202 AA; 22624 MW; 91E2716FF91353F9 CRC64;

Query Match 32.6%; Score 424.5; DB 13; Length 202;  
Best Local Similarity 45.4%; Pred.No. 3.8e-36;  
Matches 8; Conservative

QY 6 LLLLLLLVPVPLLLAAGPGGALGNRHAVYWSSNOHLRREGYTVOQNVNDLYDICYPH 65  
Db 1 LLGLLLWA-----PLLWAPPPVRRGGVYWGNSNPRELQDDYSIQVSINDHLDIYCYPH 55

QY 66 YNSSGVGAGPGGPGGAEOYVLVWSRNGRYRTCSAQ-FKRWECPNRPHAPSPKFSK 124  
Db 56 YSA-----PTPWAESTLEPMVDDEGYRGCESTPGAFKRWCKNPKFPAPVPVRFE 105

QY 125 KFORYSAFSLGYEPFHAGHEYIYSTPTHNLHWKCLRMKFVCACASTSHSGEKVP-TLPQ 18  
Db 106 KIQRFTPSLGFGRPGCTYYISVPTPGSAGRCLKLRVSVCCRASTPEPLTEVPNSOPR 165

QY 184 FTWGP 188  
Db 166 GRGGP 170

RESULT 4  
Q9PTD0 PRELIMINARY; PRT; 88 AA.

AC Q9PTD0;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DI 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE EPHRN A3 (FRAGMENT).  
OS Ctenophorus ornatus.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Iguania; Agamidae; Amphibolurinae;  
OC Ctenophorus.  
ON NCBI\_TaxID=95347;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Chen P.B., Rodger J., Dunlop S.A., Beazley L.D.;  
RL "Ephrin homologs are expressed in the adult lizard visual system.";  
RT Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
RF [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Scarborough M.T.C., Rodger J., Dunlop S.A., Beazley L.D.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF209777; AAF19444.1;  
DR InterPro; IPR001799; Ephrin.  
DR Pfam; PF00812; Ephrin; 1.  
DR ProDom; PD002533; Ephrin; 1.  
FT NON\_TER 1  
FT NON\_TER 88  
SQ SEQUENCE 88 AA; 10636 MW; C56FCD8B13F219E7 CRC64;

Query Match 31.2%; Score 406; DB 13; Length 88;  
Best Local Similarity 75.0%; Pred.No. 1.1e-34;  
Matches 72; Conservative 7; Mismatches 9; Indels 8; Gaps

QY 35 YWNSSNQHRLREGYTVOQNVNDLYDICHYNSSGVGPGAGPGGGAEQYVLVWSRN 94  
Db 1 VFWNRNPHLRREGYTVOQVSNVDLYDICHYNSS-----VPEGRVEQYILVMVNE 52

QY 95 GYRTCNASOGFKRWECNRPHAPSPKFSKEFQRY 130  
Db 53 GYTCNISOGFKRWECNRPHAPSPKFSKEFQRT 89

RESULT 5

```

Q9CZS8
ID Q9CZS8 PRELIMINARY; PRT; 206 AA.
AC Q9CZS8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 10 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2610529M21, FULL INSERT SEQUENCE.
GN EFNA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wuzhshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cdna collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012195; BAB28092.1; -.
DR MGD; MGI:106643; Efn4.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
SQ SEQUENCE 206 AA; 22859 MW; 675E32971D1C6EBC CRC64;

Query Match 31.2%; Score 406; DB 11; Length 206;
Best Local Similarity 46.9%; Pred. No. 3.3e-34;
Matches 82; Conservative 19; Mismatches 62; Indels 12; Gaps 3;

Qy 8 LLLLVVPVPLPLLAQGGGALGNRHAYVWNSNNOHLRREGYTVQVNVNDYDLYCCHYN 67
Db 3 LLPLRTVWALGLSRPLGCGSSLRHPIYWNSSNPLRLRGDAVVLGDFNDYDLYCCHYE 62
Qy 68 SSGVGPGAGPGGGAQGYLYVMSRGYRTCTNA--SQGFKRWCNRPAPHSPKPFSEK 125
Db 63 S-----PGPPEGPTFFALYVDWSEYEAETAGANSFORWNCSPFPFSPVTFSEK 114
Qy 126 FQYSFASLGEYFAGHEYYIISTPTNLHWKLCRMKVFCVC--ASTSHSGEKPV 178
Db 115 IORTPTPLGPEFLPGETYYIVISPTTSPGRCRLRLQVSCCKSGSGSHSAHPV 169

RESULT 6
O93431 PRELIMINARY; PRT; 229 AA.
AC O93431;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EPHRIN A-L1.
OS Brachydanio rerio (Zebrafish) (Zebra danio).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RT "Eph signalling is required for segmentation and differentiation of
the somites";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ006838; CAA07264.1; -.
DR InterPro; IPR001799; Ephrin.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00812; Ephrin; 1.
DR PRINTS; PR01347; EPHRIN.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN.1.
SQ SEQUENCE 229 AA; 26115 MW; 8684462F67AF6F5C CRC64;

Query Match 28.7%; Score 373; DB 13; Length 229;
Best Local Similarity 40.0%; Pred. No. 1e-30; Indels 42; Gaps 6;
Matches 84; Conservative 25; Mismatches 59;

Qy 28 ALGNRHAYVWNSNNOHLRREGYTVQVNVNDYDLYCCHYNSSGVGPGGGAQGYV 87
Db 18 ASARHSYVWNSNANFLWDDYTVDRINDYDLYDLCCHY-----AHGEIASQEAERYV 70
Qy 88 LYMSRNGYRTCTNA--SQGFKRWCNRPAPHSPKPFSEKFORYSAFSLGYEFHAGHEYY 146
Db 71 LYWLEVDYENCKPHSFQDLRWCSPFAPHAPEKSFQRTFTFLGKEFRQGESYY 130
Qy 147 ISTPTNLHWKLCRMKVFCVCSTSHSGEKPVPTLPQFTMGPNVKINVLDFEGE----- 201
Db 131 ISKPLHHGQECRLKLVDPVGVPHGSKNKK-----NVKEVEIEGKMAAGG 176
Qy 202 --NP-----QVPKLEKSI--SGTS 216
Db 177 VHNPSNRLPADDDPIAMIPVWORSVSGSGVS 206

RESULT 7
O9D7K8 PRELIMINARY; PRT; 205 AA.
ID O9D7K8;
AC O9D7K8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ADULT MALE TONGUE CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2310004J15, FULL INSERT SEQUENCE.
GN EFNA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009144; BAB26102.1; -.
DR MGD; MGI:103236; Efnal.
DR InterPro; IPR001680; WD40.
DR PRINTS; PR01347; Ephrin.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
SQ SEQUENCE 205 AA; 23772 MW; E37E55767459A4EC CRC64;

Query Match 28.2%; Score 367; DB 11; Length 205;
Best Local Similarity 38.3%; Pred. No. 3.7e-30;
Matches 85; Conservative 23; Mismatches 88; Indels 26; Gaps 6;

QY 19 PLLAAGPGALGNRAHYVWNSNOHLRREGYTVQVNVNDYDIYCPHYNSSGVGCGAGFG 78
DB 7 PLLGLCCSLAAADRHIVFWNSNPKFREEDYTVHVQLNDYDIICPHYEDDSVADAA--- 63
QY 79 PGCGAEQVLYVMSRNGYRTCN-ASQGFKEWECNRPHAPHSPIKFSKFORYSFSLGVE 137
DB 64 ----WERTLYVWEHVEYVACQPSQKQDVWNCNRPSSAKHGPEKLSVFKQFTPTPLGKE 119
QY 138 FHAGHEYIYPTTHLHWKCLRMKFVCCASTSHSGKEKVPVPLQFTMGPNVKINVLED 197
DB 120 FREGHSYVYISPIYHQSCLKLKVTV-----NG-----KITHNPQAHVNPQEK 164
QY 198 -FEGENPOVPKLEKISGTSKREHLPLAVGIAFLMTFLAS 238
DB 165 RLQADDPEVQVLSHGYSAAAPRL--FPLVWAVLLPLLLLOS 204

RESULT 8
Q9WUE7 PRELIMINARY; PRT; 102 AA.
ID Q9WUE7
AC Q9WUE7;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EPHRIN A-2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsunaga T., Davis J.G., Greene M.I.;
RT "Cellular and subcellular compartmentalization of adult peripheral
RT vestibular system by distinctive and overlapping expression of Eph
RT receptors and ephrins.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131912; AAD33515.1; -.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRODOM; PD002533; Ephrin; 1.
DR PROSITE; PS01299; EPHRIN; 1.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 11838 MW; A497302F7FD7364B CRC64;

Query Match 26.0%; Score 338; DB 11; Length 102;
Best Local Similarity 58.9%; Pred. No. 1.6e-27;
Matches 63; Conservative 13; Mismatches 23; Indels 8; Gaps 2;

QY 49 YTVQVNVNDYDIYCPHYNSSGVGCGAGPGGGAQVLYVMSRNGYRTCNASQ-GFKR 107

```

```

DB 1 YTVVEVNDYDIYCPHY-----GAPLPPAERMERYILYVNGCHASCDHRQGRFKR 53
QY 108 WECNRPHAPHSPIKFSKFORYSFSLGVEFHAGHEYIYSTPTNHL 154
DB 54 WECNRPAAPGGPLKFSKFORSLGFEFRPGHEYYVISATPPNL 100

RESULT 9
Q9PTD1 PRELIMINARY; PRT; 93 AA.
ID Q9PTD1
AC Q9PTD1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE EPHRIN A2 (FRAGMENT).
OS Ctenophorus ornatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Amphibolurinae;
OC Ctenophorus.
OX NCBI_TaxID=95347;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-BRAIN;
RA Scarbrough M.T.C., Rodger J., Dunlop S.A., Beazley L.D.;
RT "Ephrin homologs are expressed in the adult lizard visual system.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF209776; AAF19443.1; -.
DR InterPro; IPR001799; Ephrin.
DR Pfam; PF00812; Ephrin; 1.
DR PRODOM; PD002533; Ephrin; 1.
FT NON_TER 1
FT NON_TER 93
SQ SEQUENCE 93 AA; 11323 MW; BD561F18D34C0F28 CRC64;

Query Match 20.6%; Score 268.5; DB 13; Length 93;
Best Local Similarity 49.5%; Pred. No. 2.4e-20;
Matches 48; Conservative 18; Mismatches 22; Indels 9; Gaps 2;

QY 35 YVWNSNOHLRREGYTVQVNVNDYDIYCPHYNSSGVGCGAGPGGGAQVLYVMSRN 94
DB 1 VFCNRSNRFVGEYTVAVSINDYDVCYYE-----PQHSRMERYILFMVNH 52
QY 95 GYRTC-NASQGFKEWECNRPHAPHSPIKFSKFORYS 130
DB 53 GYLTCHEMRMGFKWECNRPSQPDGRLRSEKFORFT 89

RESULT 10
Q42304 PRELIMINARY; PRT; 80 AA.
ID Q42304
AC Q42304;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EPHRIN-A5 (FRAGMENT).
GN EFNA5A.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MacDonald R., Scholes J., Strahle U., Brennan C., Holder N., Brand M.,
RA Wilson S.;
RT "The Pax protein Noi protein is required for commissural axon pathway
RT formation in the rostral forebrain.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12928; CAA73391.1; -.
DR ZFIN; ZDB-GENE-001128-1; efna5a.
DR InterPro; IPR001799; Ephrin.

```



```
Db 12 ILLIFVLDTGAT-----NMPIYNSLNKRFSDKGYVLPQIGDRDLICPSSD 64
QY 68 SSGVGCAGPGGGGA--EQYVLYMVSR-----NGYRTCNASQGFKRWEK 112
Db 65 P-----PGPRADYEVYKLYLVSSREQADRCVTCAPNLLTCD----- 104
QY 113 PHAPHSPIKFESEKFORYSAFSLGYEFHAGHEYIYSTPHNLHMKCLRMKVFVCCASTSH 172
Db 105 --KPNSDMRFTIKFOYSNLMGHEFKTNHDYFIAT-SDGTROGLSMRGVCAT---- 157
QY 173 SGEKVPPTLPQFTMGPNVK-----INVLEDFEGE--NPQVPKLEKSISGTSKPREH 221
Db 158 QGMKVVLVKQSPYGLPAKSPKPDAGRNPNPGTGNSTHPQIIP-----RGSGENG 212
QY 222 LP-----LAVGTAFLLM 233
Db 213 LPASNIATVIAAGAGGSAFLLL 233

RESULT 14
ID Q9PT69 PRELIMINARY; PRT: 327 AA.
AC Q9PT69;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE EPHRIN-B3 PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE=20099673; PubMed=10633856;
RA Helbling P.M., Saulnier D.M.E., Robinson V., Christiansen J.H.,
RA Wilkinson D.G., Brandli A.W.;
RT "Comparative analysis of embryonic gene expression defines potential
RT interaction sites for Xenopus EphA4 receptors with ephrin-B ligands.";
RL Dev. Dyn. 216:361-373(1999).
DR EMBL: AJ236866; CAB65511.1.
DR InterPro: IPR002086; Aldohyde_dehydr.
DR InterPro: IPR001799; Ephrin.
DR Pfam: PF00812; Ephrin; 1.
DR ProDom: PD002533; Ephrin; 1.
DR ProSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE: PS01299; EPHRIN; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 327 AA; 35913 MW; 4BB0FA39D4C22DCD CRC64;

Query Match 12.6%; Score 164.5; DB 13; Length 327;
Best Local Similarity 26.6%; Pred. No. 8e-09;
Matches 63; Conservative 31; Mismatches 78; Indels 65; Gaps 13;

QY 35 VYWNSSNQHRL-REGYTVQVNVNDYDIYCPHYNSGVGPGAGPGGABQYVLYMV-- 91
Db 30 IYWNSSNKRFEDETEGYLVLPQIGDRDLCLCPSEPPQ-----GPFSSPYEYKLYLVGT 83
QY 92 -----SRNGYRTCNASQGFKRWEKRNHAPHSPKFESEKFORYSAFSLGYEFHAGHEYI 146
Db 84 KEEMSCSILATPNL-----LLTCDR---PSQDLRTIKFOEFSPLNLWGHEFQSDRYI 135
QY 147 ISTPHNL-----HWKCLRMKVFVCCASTSHSGEKPVPPTLPQFT-----MGPNV 190
Db 136 IATSDGTMDGIETLQGGVCTKGMKVLKV---GQSPNGATP-PRRPSSACKDSGISPSV 191
QY 191 KINVLDFEGENPQVPKLEKSISGTSKPR-----EHLPL-----AVGTAFLLMTF 235
Db 192 -----PNPDIPNVGET-SGNATKGTENGPLPISHVPLVAGAGGAALLLVF 237
```

```
RESULT 15
ID Q9PUJ4 PRELIMINARY; PRT: 333 AA.
AC Q9PUJ4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE EPHRIN-B2 PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Menzel P., Pasquale E.B.;
RT "Coding sequence of chicken ephrin-B2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF180729; AAD53948.1;
DR InterPro: IPR001799; Ephrin.
DR Pfam: PF00812; Ephrin; 1.
DR ProDom: PD002533; Ephrin; 1.
DR PROSITE: PS01299; EPHRIN; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 333 AA; 36761 MW; 4C28E8CB211B7783 CRC64;

Query Match 12.5%; Score 162.5; DB 13; Length 333;
Best Local Similarity 27.1%; Pred. No. 1.3e-08;
Matches 61; Conservative 29; Mismatches 92; Indels 43; Gaps 9;

QY 35 VYWNSSN-QHLRREGYTVQVNVNDYDIYCPHYNSGVGPGAGPGGABQYVLYMVSR 93
Db 35 IYWNSSNPKFLPGGGLVLPQIGDKLDIICPKVDSKTVGQ-----VEYKVMWVK 85
QY 94 NGYTCNASQGFKRWEKRNHAPHSPKFESEKFORYSAFSLGYEFHAGHEYIYST 149
Db 86 DOADSALRKD-----NTPLLNCAPQDQVVKFTIKFOEFSPLNLWGHEFQKNDYIYST 139
QY 150 PTHNLH-----WKCLRMKVFVCCASTSHSGEKPVPPT-----LPQFTMGPNVKIN-- 194
Db 140 SNGSLEGLNNQGGVCTKTKMLKMKVQDPNSAGLPRSTDPTKRPRQEAGTNGKSTTS 199
QY 195 --LEDFEGENPQVKL-EKISGTSKPREHPLAVGTAFLLMTPL 236
Db 200 PFVKDHSSTGDSKAGHSILGS-----EVALFAGIASGCIIFI 239
```

Search completed: July 13, 2002, 10:52:33  
Job time: 463 sec



```
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 71 74 MISSING (IN REF. 2)
SQ SEQUENCE 238 AA; 26350 MW; 8EFD0A8FE33FDDA CRC64;

Query Match 100.0%; Score 1301; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLPLLLLLLVPLPLLAQGGGALGNHAYVWSSNOHLRREGYTVQVNVNDYLD 60
DB 1 MAAPLPLLLLLLVPLPLLAQGGGALGNHAYVWSSNOHLRREGYTVQVNVNDYLD 60

QY 61 IYCPHYNSGVPAGPGGAGAEYVLYVMSRNGYRTCNASQGFKRWCNRPAPHSP 120
DB 61 IYCPHYNSGVPAGPGGAGAEYVLYVMSRNGYRTCNASQGFKRWCNRPAPHSP 120

QY 121 KFSEKFORYSAFSLGYEFHAGHEYYIYSTPTNHLHWKCLRMKVFCVCCASTSHSGEKP 180
DB 121 KFSEKFORYSAFSLGYEFHAGHEYYIYSTPTNHLHWKCLRMKVFCVCCASTSHSGEKP 180

QY 181 LPQFTMGPNVKINLVDFEGENPOVPKLEKSISCTSPKREHLPLAVGIAFFLMTFLAS 238
DB 181 LPQFTMGPNVKINLVDFEGENPOVPKLEKSISCTSPKREHLPLAVGIAFFLMTFLAS 238

RESULT 2
EFA3_MOUSE STANDARD; PRT; 187 AA.
AC 008545; 055217;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Ephrin-A3 (EPH-related receptor tyrosine kinase ligand 3) (LERK-3)
DE (EHK1 ligand) (EHK1-L) (Fragment).
DE EFNA3 OR EPLG3 OR LERK3 OR EPL3.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX Carretti D.P., Nelson N.;
MEDLINE=98126446; PubMed=9465306;
RA "Characterization of the genes for mouse LERK-3/Ephrin-A3 (Epl3),
RT mouse LERK-4/Ephrin-A4 (Epl4), and human LERK-6/Ephrin-A2 (EPLG6):
RT conservation of intron/exon structure.";
RL Genomics 47:131-135(1998).
RN [2]
RP SEQUENCE OF 17-78 FROM N.A.
RC MEDLINE=97060319; PubMed=8903354;
RA Flenniken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;
RT "Distinct and overlapping expression patterns of ligands for
RT Eph-related receptor tyrosine kinases during mouse embryogenesis.";
RL Dev. Biol. 179:382-401(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U92885; AAC39961.1; -
CC EMBL; U90666; AAB50241.1; -
CC MGD; MGI:106644; Efn3.
CC InterPro; IPR001799; Ephrin.
```

```
DR Pfam: PF00812; Ephrin; 1.
DR ProDom: PD002533; Ephrin; 1.
DR PROSITE; PS01293; EPHRIN; 1.
KW Glycoprotein; GPI-anchor.
FT NON_TER 1 1
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 19 19 Y -> I (IN REF. 2).
FT CONFLICT 33 34 EQ -> DR (IN REF. 2).
FT CONFLICT 46 47 RT -> QP (IN REF. 2).
FT CONFLICT 78 78 Y -> W (IN REF. 2).
SQ SEQUENCE 187 AA; 21171 MW; CCE4915751760743 CRC64;

Query Match 74.0%; Score 963; DB 1; Length 187;
Best Local Similarity 91.3%; Pred. No. 1.3e-83;
Matches 178; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 44 LRREGYTVQVNVNDYDIYCPHYNSGVPAGPGGAGAEYVLYVMSRNGYRTCNASQ 103
DB 1 MRREGYTVQVNVNDYDIYCPHYNSGVPAGPGGAGAEYVLYVMSRNGYRTCNASQ 52

QY 104 GFRWCNRPAPHSPKIFSEKFORYSAFSLGYEFHAGHEYYIYSTPTNHLHWKCLRMKV 163
DB 53 GSKRWECNRHSHSPKIFSEKFORYSAFSLGYEFHAGHEYYIYSTPTNHLHWKCLRMKV 112

QY 164 FVCCASTSHSGEKPVTLPQFTMGPNVKINLVDFEGENPOVPKLEKSISCTSPKREHLP 223
DB 113 FVCCASTSHSGEKPVTLPQFTMGPNVKINLVDFEGENPOVPKLEKSISCTSPKREHLP 172

QY 224 LAVGIAFFLMTFLAS 238
DB 173 LAVGIAFFLMTLLAS 187

RESULT 3
EFA2_MOUSE STANDARD; PRT; 209 AA.
AC P52801;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A2 precursor (EPH-related receptor tyrosine kinase ligand 6)
DE (LERK-6) (ELF-1) (CEK7-ligand) (CEK7-L).
DE EFNA2 OR EPLG6 OR LERK6 OR ELF1 OR EPL6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER; TISSUE=Brain;
RX MEDLINE=95007776; PubMed=7522971;
RA Cheng H.J., Flanagan J.G.;
RT "Identification and cloning of ELF-1, a developmentally expressed
RT ligand for the Mek4 and Sek receptor tyrosine kinases.";
RL Cell 79:157-168(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95181289; PubMed=7876076;
RA Shao H., Lou L., Pandey A., Verderame M.F., Siever D.A., Dixit V.M.;
RT "cDNA cloning and characterization of a Cek7 receptor
RT protein-tyrosine kinase ligand that is identical to the ligand
RT (ELF-1) for the Mek-4 and Sek receptor protein-tyrosine kinases.";
RL J. Biol. Chem. 270:3467-3470(1995).
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA3, EPHA4 AND
CC EPHA5.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
```



[illegible]

Dd	173 T	173
<hr/>		
RESULT 7		
EFA5_HUMAN	STANDARD;	PRT; 228 AA.
ID EFA5_MOUSE		
AC P52803;		
DT 01-OCT-1996	(Rel. 34, Created)	
DT 01-OCT-1996	(Rel. 34, Last sequence update)	
DT 16-OCT-2001	(Rel. 40, Last annotation update)	
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)		
DE (LERK-7) (AL-1).		
GN EFNA5 OR EPLG7 OR LERK7.		
OS Homo sapiens (Human).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX NCBI_TaxID=9606;		
[1]		
RN SEQUENCE FROM N.A.		
RP MEDLINE=95267434; PubMed=7748564;		
RA Winslow J.W., Moran P., Valverde J., Shih A., Yuan J.O., Wong S.C.,		
RA Tsai S.P., Goddard A., Henzel W.J., Heftl F.;		
RT Cloning of AL-1, a ligand for an Eph-related tyrosine kinase		
RT receptor involved in axon bundle formation.";		
RL Neuron 14:973-981(1995).		
[2]		
RN SEQUENCE FROM N.A.		
RP TISSUE=Brain;		
RC MEDLINE=9732664; PubMed=9245480;		
EX Kozlosky C.J., Vandenbos T., Park L.S., Cerretti D.P., Carpenter M.K.;		
RT "LERK-7: a ligand of the Eph-related kinases is developmentally		
RT regulated in the brain.";		
RL Cytokine 9:540-549(1997).		
[3]		
RN FUNCTION		
RP MEDLINE=20069483; PubMed=10601038;		
EX Davy A., Gale N.W., Murray E.W., Klinghoffer R.A., Soriano P.,		
RA Feuerstein C., Robbins S.M.;		
RA *Compartmentalized signaling by GPI-anchored ephrin-A5 requires the		
RT Fyn tyrosine kinase to regulate cellular adhesion.";		
RL Genes Dev. 13:3125-3135(1999).		
[4]		
CC -!- FUNCTION: MAY FUNCTION ACTIVELY TO STIMULATE AXON FASCICULATION.		
CC INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE		
CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS		
CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF		
CC THE FYN TYROSINE KINASE.		
CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND		
CC EPHB1.		
CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE		
CC MICRODOMAINS.		
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC or send an email to license@isb-sib.ch).		
CC [1]		
DR EMBL; U26403; AAB60377.1; --		
DR MIM; 601535; --		
DR InterPro: IPR001799; Ephrin.		
DR Pfam: PF00812; Ephrin; 1.		
DR PRINTS; PR01347; EPHRIN.		
DR ProDom; PD002533; Ephrin; 1.		
DR PROSITE; PS01299; EPHRIN; 1.		
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal;		
KW Polymorphism.		
SIGNAL 1 20 POTENTIAL.		
FT CHAIN 21 228 EPHRIN-A5.		
N-CARBOHYD 37 37 N-LINKED (GLCNAC... ) (POTENTIAL).		

```

DR EMBL: U90664; AAB50239.1; -
DR EMBL: U90665; AAB50240.1; -
DR MGI: 107444; EfnA5.
DR InterPro: IPR001799; Ephrin.
DR Pfam: PF00812; Ephrin; 1.
DR PRINTS: PR01347; EPHRIN.
DR PRODOM: PD002533; Ephrin; 1.
DR PROSITE: PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal;
KW Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 228 EPHRIN-A5.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLOC 163 189 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 228 AA; 26339 MW; 85439F537420022 CRC64;

Query Match 31.9%; Score 414.5; DB 1; Length 228;
Best Local Similarity 39.5%; Pred. No. 7.1e-32;
Matches 98; Conservative 36; Mismatches 67; Indels 47; Gaps 11;

QY 8 LLLLLVPLPLLAQPGG-ALGNRHAVYWNSSNHLRREGYTVQVNVNDYLDIYCPHY 66
   :||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 MLTLFLVLMWCVFSQDPGKVVADRYAVYWNSSNPRQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGGGAQYVLYVMSRNGYRTC-NASQGFKEWECNRPHAPHSPIKFSEK 125
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 EDS-----VPEDKTERVYLVYVNFQDYSACDHTSGKFKWECNRPHSPNGLKFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYVYISTP-THNLHWKCLRMKVFV-----CCASTSHSGEKVPPT 180
   || : : ||||| : || : ||||| : || : ||||| : || : ||||| : || :
Db 118 FQFTFPGSLGFEPGRGYFYISSAIPDNGRRSCLKLKVFVRYPTNSCMKTCGVDRD--- 173

QY 181 LPQFTMGPNVKINVLDFE-----GEN-PQVPKLEKISGTSKREHLPLAVGI 228
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 --VFDVNDKVE-NSLEPADDTVHESAEPSSRGENAAQTPI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236
   || : :
Db 218 LLFLLAML 225

RESULT 9
EFAS_RAT
ID EFAS_RAT STANDARD; PRT; 228 AA.
AC P97605;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)
DE (LERK-7) (AL-1).
GN EFNA5 OR EPLG7 OR LERK7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC STRAIN=SPRAGUE-DAWLEY.
RX MEDLINE=95267434; PubMed=7748564;
RA Winslow J.W., Moran P., Valverde J., Shih A., Yuan J.Q., Wong S.C.,
RA Tsai S.P., Goddard A., Henzel W.J., Hefti F., Beck K.D., Caras I.W.;
RT "Cloning of AL-1, a ligand for an Eph-related tyrosine kinase
RT receptor involved in axon bundle formation.";
RL Neuron 14:973-981(1995).
RN [2]
RC STRAIN=SPRAGUE-DAWLEY.
RA Li Y.Y., McTiernan C.F., Feldman A.M.;
RT "LERK7, rat ligand for Eph-related receptor tyrosine kinase.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- INDUCES COMPARTMENTALIZED SIGNALING WITHIN A CAVEOLAE-LIKE

```

```

CC MEMBRANE MICRODOMAIN WHEN BOUND TO THE EXTRACELLULAR DOMAIN OF ITS
CC COGNATE RECEPTOR. THIS SIGNALING EVENT REQUIRES THE ACTIVITY OF
CC THE EFN TYROSINE KINASE (BY SIMILARITY).
CC -!- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHA2, EPHA3 AND
CC EPHB1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR IT
CC IS COMPARTMENTALIZED IN DISCRETE CAVEOLAE-LIKE MEMBRANE
CC MICRODOMAINS (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART, PLACENTA AND LUNG.
CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U69279; AAC05801.1; -
CC InterPro: IPR001799; Ephrin.
CC Pfam: PF00812; Ephrin; 1.
CC PRINTS: PR01347; EPHRIN.
CC PRODOM: PD002533; Ephrin; 1.
CC PROSITE: PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 228 EPHRIN-A5.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 228 AA; 26358 MW; 855985532D580022 CRC64;

Query Match 31.9%; Score 414.5; DB 1; Length 228;
Best Local Similarity 39.5%; Pred. No. 7.1e-32;
Matches 98; Conservative 36; Mismatches 67; Indels 47; Gaps 11;

QY 8 LLLLLVPLPLLAQPGG-ALGNRHAVYWNSSNHLRREGYTVQVNVNDYLDIYCPHY 66
   :||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 MLTLFLVLMWCVFSQDPGKVVADRYAVYWNSSNPRQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGGGAQYVLYVMSRNGYRTC-NASQGFKEWECNRPHAPHSPIKFSEK 125
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 EDS-----VPEDKTERVYLVYVNFQDYSACDHTSGKFKWECNRPHSPNGLKFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYVYISTP-THNLHWKCLRMKVFV-----CCASTSHSGEKVPPT 180
   || : : ||||| : || : ||||| : || : ||||| : || : ||||| : || :
Db 118 FQFTFPGSLGFEPGRGYFYISSAIPDNGRRSCLKLKVFVRYPTNSCMKTCGVDRD--- 173

QY 181 LPQFTMGPNVKINVLDFE-----GEN-PQVPKLEKISGTSKREHLPLAVGI 228
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 --VFDVNDKVE-NSLEPADDTVHESAEPSSRGENAAQTPI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236
   || : :
Db 218 LLFLLAML 225

RESULT 10
EF4_MOUSE
ID EF4_MOUSE STANDARD; PRT; 206 AA.
AC O08542; O55218;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-A4 precursor (EPH-related receptor tyrosine kinase ligand 4)
DE (LERK-4).
GN EFNA4 OR EPLG4 OR LERK4 OR EPL4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.

```

RX MEDLINE=97060319; PubMed=8903354;  
 RA Flenniken A.M., Gale N.W., Yancopoulos G.D., Wilkinson D.G.;  
 RT "Distinct and overlapping expression patterns of ligands for  
 RT Eph-related receptor tyrosine kinases during mouse embryogenesis";  
 RL Dev. Biol. 179:382-401(1996).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=129;  
 RX MEDLINE=98126446; PubMed=9465306;  
 RA Cerretti D.P., Nelson N.;  
 RT "Characterization of the genes for mouse LERK-3/Ephrin-A3 (Epl3),  
 RT mouse LERK-4/Ephrin-A4 (Epl4), and human LERK-6/Ephrin-A2 (EPLG6):  
 RT conservation of intron/exon structure.";  
 RL Genomics 47:131-135(1998).  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (BY  
 CC similarity).  
 CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U90663; AAC50238.1; -;  
 DR EMBL; U92890; AAC39962.1; -;  
 DR EMBL; U92889; AAC39962.1; JOINED.  
 DR MGI; 106643; Efn4.  
 DR InterPro: IPR001799; Ephrin.  
 DR Pfam; PF00812; Ephrin; 1.  
 DR PRINTS; PR01347; Ephrin.  
 DR PRODOM; PD002533; Ephrin; 1.  
 DR PROSITE; PS01299; EPHRIN; 1.  
 KW Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 206 EPHRIN-A4.  
 FT CARBOHYD 33 33 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT SITE 41 43 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CONFLICT 1 4 MRL -> MLLRLGLVPTPTPPGGLV (IN REF.  
 FT 1).  
 SQ SEQUENCE 206 AA; 22861 MW; 43501971DDIC6EA5 CRC64;  
  
 Query Match 31.5%; Score 410; DB 1; Length 206;  
 Best Local Similarity 47.4%; Pred. No. 1.7e-31;  
 Matches 83; Conservative 18; Mismatches 62; Indels 12; Gaps 3;  
  
 Qy 8 LLLLVVPLLLPALLAOGPGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLDIYCPHYN 67  
 Db 3 LPLLVTVWALLGSLRPLGSSLRHPYWNSSNPLRLRGDAVVELGFDYLDIFCPHYE 62  
 Qy 68 SSGVGPGAGPGGGAGQYVLYMSRNGYRTCA--SQGFKRWECNRPAPHSPKPFSEK 125  
 Db 63 S-----PPGPGPFTFALLYVMDWSGEACTAEGANAFQWNCNMPFAPSPVRESEK 114  
 Qy 126 FQYSFASLGYEPHAGHEYIYIPTNHLNHLKRLMKVFCVCC--ASTSHSGEKPV 178  
 Db 115 IQRTVPLPLGLFEFLPGETIYIYIYPTPEPSPGRLRLQVSVCCRESGSSHSAHPV 169  
  
 RESULT 11  
 ID EFA5\_CHICK STANDARD; PRT; 228 AA.  
 AC P52804;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)  
 DE (LERK-7) (RAGS protein).  
 GN EFNA5 OR RAGS.

OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Posterior tectum;  
 RX MEDLINE=95360980; PubMed=7634326;  
 RA Drescher U., Kremser C., Handwerker C., Loschinger J., Noda M.,  
 RA Bonhoeffer F.;  
 RT "In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa  
 RT tectal protein related to ligands for Eph receptor tyrosine  
 RT kinases";  
 RL Cell 82:359-370(1995).  
 CC -!- FUNCTION: INDUCES GROWTH CONE COLLAPSE AND REPULSION OF RETINAL  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
 CC (Potential).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN A GRADED FASHION ACROSS THE  
 CC TECTUM BEING MORE STRONGLY EXPRESSED TOWARDS THE POSTERIOR POLE.  
 CC -!- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X90377; CAA62027.1; -;  
 DR InterPro; IPR001799; Ephrin.  
 DR Pfam; PF00812; Ephrin; 1.  
 DR PRINTS; PR01347; EPHRIN.  
 DR PRODOM; PD002533; Ephrin; 1.  
 DR PROSITE; PS01299; EPHRIN; 1.  
 KW Developmental protein; Neurogenesis; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 228 EPHRIN-A5.  
 FT CARBOHYD 37 37 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 228 AA; 26206 MW; 56D8E4FBDECF18AD CRC64;  
  
 Query Match 31.1%; Score 404.5; DB 1; Length 228;  
 Best Local Similarity 49.1%; Pred. No. 6.2e-31;  
 Matches 79; Conservative 29; Mismatches 42; Indels 11; Gaps 4;  
  
 Qy 8 LLLLVVPLLLPALLAOGPGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLDIYCPHY 66  
 Db 6 MLLLVAALWVCRGQEPGRKAVADRYAVYWNSTNPRFQGDYHIDVICINDYLDVFCPHY 65  
 Qy 67 NSSGVGPGAGPGGGAGQYVLYMSRNGYRTCA--NASQGFKRWECNRPAPHSPKPFSEK 125  
 Db 66 EDS-----VPEDKTERVLYVMVNDGYSSCDHISKGFKRWECNRPHPSPGLKPFSEK 117  
 Qy 126 FQYSFASLGYEPHAGHEYIYIPTNHLNHLKRLMKVFPV 165  
 Db 118 FQLTFTPLGLFEFLPGREYFYIISAITPDNGRRSCLKLKVEV 158  
  
 RESULT 12  
 ID EFA5\_BRARE STANDARD; PRT; 228 AA.  
 AC P79728;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Ephrin-A5 precursor (EPH-related receptor tyrosine kinase ligand 7)  
 DE (LERK-7) (AL-1) (2FEPLH4).  
 GN EFNA5 OR EFNA5B OR AL1 OR EPLG7 OR LERK7.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





FT CONFLICT 204 204 S -> T (IN REF. 1).  
SQ SEQUENCE 205 AA; 23802 MW; 5A8F3A6E2091E868 CRC64;

Query Match 28.7%; Score 374; DB 1; Length 205;  
Best Local Similarity 38.7%; Pred. No. 4e-28;  
Matches 86; Conservative 23; Mismatches 87; Indels 26; Gaps 6;  
QY 19 PLLAQGGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLDIYCPHYNSSGVGPGAGPG 78  
Db 7 PLLGLCCSLAAADRHIVFNWSSNPKFREEDYTVHVQLNDYLDIICPHYEDDSVADAA--- 63  
QY 79 PGCGAEQYVLYMYSRNGYRTCN-ASOGFKRWEKNRPHAPHSPIKFSEKFORYSAFSLGYE 137  
Db 64 ----MERYTLYMVEHQEYVACQPQSKDQVRWNCNRPNSAKHGPEKLFSEKFORFTPFILGKE 119  
QY 138 FHAGHEYVYISTPTNHLHWKCLRMKVFCVCCASTSHSGEKPVPPTLPQFTMGPNVKINVL 197  
Db 120 FKEGHSYVYISKPIYHQESQCLKLVTV-----KITHNPOAHVNPQEK 164  
QY 198 -FEGENPOVPKLEKSIKSGTSPKREHLPLAVGIAFFLMTFLAS 238  
Db 165 RLOADDPEQVQLHSICGSAAPRL--FPLVWAVLLPLLLLOS 204

Search completed: July 13, 2002, 10:53:18  
Job time: 463 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:55:25 ; Search time 42.58 Seconds  
(without alignments)  
537.089 Million cell updates/sec

Title: US-09-733-756-2  
Perfect score: 1301  
Sequence: 1 MAAAPLLLLLLVPVLLPL.....REHLPLAVGIAFFLMTFLAS 238  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1301	100.0	238	2 I38849	LERK-3 - human
2	437	33.6	209	2 A54984	ELF-1 protein precursor
3	434	33.4	213	2 JE0322	ephrin-A2 - human
4	415.5	31.9	228	2 I58170	LERK-7 precursor
5	404.5	31.1	228	2 A57084	repulsive axon guidance
6	390	30.0	205	2 A36377	B61 protein precursor
7	364.5	28.0	201	2 I38850	LERK-4 - human
8	171.5	13.2	333	2 I84743	hepatoma transmembrane
9	171.5	13.2	336	2 I49766	hepatoma transmembrane
10	148.5	11.4	345	2 S46993	elk ligand - human
11	143.5	11.0	462	2 T32645	hypothetical protein
12	139.5	10.7	345	2 I48780	Stral/Eplg2 protein
13	132.5	10.2	345	2 I58406	LERK-2 - rat
14	100	7.7	356	2 T40265	hypothetical zinc-finger
15	91	7.0	2206	1 GNV27	genome polyprotein
16	88.5	6.8	1024	2 E90678	beta-D-galactosidase
17	88.5	6.8	1024	2 A85529	beta-D-galactosidase
18	87	6.7	817	2 T21336	hypothetical protein
19	85.5	6.6	709	2 T28712	hypothetical protein
20	84	6.5	1027	2 T19173	hypothetical protein
21	83.5	6.4	488	2 S67744	hypothetical protein
22	82.5	6.3	432	2 A25483	env polyprotein, r
23	82.5	6.3	466	1 TWFF	transcription factor
24	82	6.3	237	2 T19914	hypothetical protein
25	82	6.3	334	2 T16772	hypothetical protein
26	82	6.3	376	2 B84463	hypothetical protein
27	82	6.3	728	2 A48830	probable transmembrane
28	82	6.3	1144	2 A75132	hypothetical protein
29	81.5	6.3	1173	2 I50620	prockr2 - chicken

30 81 6.2 457 2 T19109 hypothetical prote  
31 81 6.2 1613 2 S39059 protein BRG1 - hum  
32 81 6.2 1647 2 S45252 SNF2beta protein -  
33 79.5 6.1 434 2 B72410 alkaline phosphata  
34 79 6.1 612 2 A54282 reversed polarity  
35 79 6.1 859 1 VCLJMN env polyprotein pr  
36 78.5 6.0 202 2 T01605 phytocyanin At2g44  
37 78.5 6.0 498 2 B84789 probable protein w  
38 78.5 6.0 1024 1 GBEC beta-galactosidase  
39 78.5 6.0 1291 2 T13389 hypothetical prote  
40 78 6.0 321 2 I38238 transcription fact  
41 78 6.0 328 1 S05426 chitinase (EC 3.2.  
42 78 6.0 700 2 A32392 protein kinase C.  
43 78 6.0 2206 1 GNNY4P genome polyprotein  
44 78 6.0 2206 2 S03822 genome polyprotein  
45 77.5 6.0 570 2 H90370 thermopsin precurs

ALIGNMENTS

RESULT 1  
I38849  
LERK-3 - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 29-Sep-1999  
C:Accession: I38849  
R:Kozlosky, C.J.; Maraskovsky, E.; McGrew, J.T.; VandenBos, T.; Teepe, M.; Lyman, S.  
Oncogene 10, 299-306, 1995  
A:Title: Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs e  
A:Reference number: I38849; MUID:95140419  
A:Accession: I38849  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-238 <RES>  
A:Cross-references: EMBL:U14187; NID:g642832; PIDN:AAC50078.1; PID:g642833  
C:Genetics:  
A:Gene: GDB:EPLG3  
A:Cross-references: GDB:438336; OMIM:601381  
A:Map position: lq21-lq22  
C:Superfamily: axon guidance signal protein

Query Match 100.0%; Score 1301; DB 2; Length 238;  
Best Local Similarity 100.0%; Pred. No. 1.7e-114; Indels 0; Gaps 0;  
Matches 238; Conservative 0; Mismatches 0;  
Qy 1 MAAAPLLLLLLVPVLLPLLAQGGALGNHRAVYNNSSNQHLRRREGYTVQVNVNDYLD 60  
Db 1 MAAAPLLLLLLVPVLLPLLAQGGALGNHRAVYNNSSNQHLRRREGYTVQVNVNDYLD 60  
Qy 61 IYCPHYSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCTNASQGFKECRPHAPHSPI 120  
Db 61 IYCPHYSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCTNASQGFKECRPHAPHSPI 120  
Qy 121 KFSEKFORSAFSGYEFHAGHEYYIYSTPTNHLHWKLMKVFVCCASTSHSGEKVPPT 180  
Db 121 KFSEKFORSAFSGYEFHAGHEYYIYSTPTNHLHWKLMKVFVCCASTSHSGEKVPPT 180  
Qy 181 LPQFTMGPNVKINVLDFEGENPOVKLEKSIGTSKPRHLPLAVGIAFFLMTFLAS 238  
Db 181 LPQFTMGPNVKINVLDFEGENPOVKLEKSIGTSKPRHLPLAVGIAFFLMTFLAS 238

RESULT 2  
A54984  
ELF-1 protein precursor - mouse  
N:Alternate names: Cek7 ligand  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 29-Sep-1999  
C:Accession: A54984; A55873  
R:Cheng, H.J.; Flanagan, J.G.  
Cell 79, 157-168, 1994

A:Title: Identification and cloning of ELF-1, a developmentally expressed ligand for the  
A:Reference number: A54984; MUID:95007776  
A:Accession: A54984  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-209 <CHE>  
A:Cross-references: GB:U14941; NID:9558836; PIDN:AAA53636.1; PID:9558837  
R:Shao, H.; Lou, L.; Pandey, A.; Verderame, M.F.; Siever, D.A.; Dixit, V.M.  
J. Biol. Chem. 270, 3467-3470, 1995  
A:Title: cDNA cloning and characterization of a Cdk7 receptor protein-tyrosine kinase 11  
A:Reference number: A55873; MUID:95181289  
A:Accession: A55873  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-209 <SHA>  
A:Cross-references: GB:U14752; NID:9681886; PIDN:AAA68520.1; PID:9681887  
C:Superfamily: axon guidance signal protein  
C:Keywords: lipoprotein; membrane protein

Query Match 33.6%; Score 437; DB 2; Length 209;  
Best Local Similarity 50.5%; Pred. No. 1.6e-33;  
Matches 96; Conservative 25; Mismatches 47; Indels 22; Gaps 7;

QY 2 AAAPLLLLLVPPVPLLLA--QGGGALGNRHAVYWNSSNQHLRRE-----GYTVQV 53  
DB 4 AQRPPLPLLLL-----LLPLRARNEDPARANADRYAVYWNRSNPRFOVSAGDGGGYTVEV 59

QY 54 NVNDYLDIYCPHYNSGVGPGGAGPGGAGQYVLYWVSRNGYRTCNASQ-GFKRWECNR 112  
DB 60 SINDYLDIYCPHY-----GAPLPPAERMERYILYVWNGEGHASCDDRGRGFKRWECNR 112

QY 113 PHAPHSPIKSEKFORYSAFSLGFEFHAGHEHYIIS-TPTHNLHWKCLRMKVFCVCCAS-T 170  
DB 113 PAAPGGPLAFSEKQFLTFPSLGFEPGHEHYIISATPPNVLVDRPCLRLKLVYVRPTNET 172

QY 171 SHSGEKPVP 180  
DB 173 LYEAPPEIPT 182

RESULT 3  
JE0322  
eprIn-A2 - human  
C:Species: Homo sapiens (man)  
C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: JE0322  
R:Aasheim, H.; Pedetour, F.; Grosgeorge, J.; Logtenberg, T.  
Biochem. Biophys. Res. Commun. 252, 378-382, 1998  
A:Title: Cloning, chromosomal mapping, and tissue expression of the gene encoding the huma  
A:Reference number: JE0322; MUID:95045414  
A:Accession: JE0322  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-213 <AAS>  
A:Cross-references: GB:AJ007292; NID:93688367; PIDN:CAA07435.1; PID:93688368  
C:Superfamily: axon guidance signal protein

Query Match 33.4%; Score 434; DB 2; Length 213;  
Best Local Similarity 50.5%; Pred. No. 3.2e-33;  
Matches 96; Conservative 24; Mismatches 52; Indels 18; Gaps 7;

QY 2 AAAPLL-LLLLLVPPVPLLLA--AQPGGALGNRHAVYWNSSNQHLR-----EGYTVQV 53  
DB 4 AQAPLLPLLLLLPLPPPPAPPDRRRANSRDYAVYWNRSNPRFHAGAGDGGGYTVEV 63

QY 54 NVNDYLDIYCPHYNSGVGPGGAGQYVLYWVSRNGYRTCNASQ-GFKRWECNR 112  
DB 64 SINDYLDIYCPHY-----GAPLPPAERMERYILYVWNGEGHASCDDRGRGFKRWECNR 116

QY 113 PHAPHSPIKSEKFORYSAFSLGFEFHAGHEHYIIS-TPTHNLHWKCLRMKVFCVCCAS-T 170  
DB 113 PAAPGGPLAFSEKQFLTFPSLGFEPGHEHYIISATPPNVLVDRPCLRLKLVYVRPTNET 172

DB 117 PAAPGGPLAFSEKQFLTFPSLGFEPGHEHYIISATPPNVLVDRPCLRLKLVYVRPTNET 176  
QY 171 SHSGEKPVP 180  
DB 177 LYEAPPEIPT 186

RESULT 4  
I58170  
LERK-7 precursor - human  
N:Alternate names: AL-1  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
C:Accession: I58170; G01812  
R:Winslow, J.W.; Moran, P.; Valverde, J.; Shih, A.; Yuan, J.Q.; Wong, S.C.; Tsai, S.P.  
Neuron 14, 973-981, 1995  
A:Title: Cloning of AL-1, a ligand for an Eph-related tyrosine kinase receptor involv  
A:Reference number: I58170; MUID:95267434  
A:Accession: I58170  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-228 <RES>  
A:Cross-references: GB:S77167; NID:9914184; PID:9914185  
R:Kozlosky, C.J.; Vandenbos, T.; Park, L.S.; Cerretti, D.P.; Carpenter, M.K.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: G08477  
A:Accession: G01812  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-228 <KOZ>  
A:Cross-references: EMBL:U26403; NID:g1019430; PIDN:AAB60377.1; PID:g1019431  
C:Genetics:  
A:Gene: GDB:EPUG7; AFI: LERK7  
A:Cross-references: GDB:568757; OMIM:601535  
A:Map position: 13q33-13q33  
C:Superfamily: axon guidance signal protein

Query Match 31.9%; Score 415.5; DB 2; Length 228;  
Best Local Similarity 39.5%; Pred. No. 1.9e-31;  
Matches 98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

QY 8 LLLLVPPVPLLLAQQPGG-ALGNRHAVYWNSSNQHLRREGYTVQVNVNDYLDIYCPHY 66  
DB 6 MLTLVFLVLMCVSPQDPGSKAVADRYAVYWNSSNPRFQGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGGGAGQYVLYWVSRNGYRTCNASQGFKRWECNRPHAPHSPIKFSK 125  
DB 66 EDS-----VPEDKTERVLYVWVNFEDGYSACDHTSGFKRWECNRPHSPNGPLKFSK 117

QY 126 FQYSASFSLGVEFHAGHEHYIIS-TPTHNLHWKCLRMKVVF---CCASTSHSGEKPVP 180  
DB 118 FQLTFFSLGFEFRPGREYFISSAIPDNGRRSCLKLVFVRPTNSCMKTIIGVHIDR---- 173

QY 181 LPQTMGPNVKINLVDEP-----GEN-POVPKLEKISGTSKREHLPLAVGI 228  
DB 174 --VFDVNDKVE-NSLEPADDTVHESAPSRGENAAQTPI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236  
DB 218 LLFLML 225

RESULT 5  
A57084  
repulsive axon guidance signal protein RAGS precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Jun-2000  
C:Accession: A57084  
R:Drescher, U.; Krenoser, C.; Handwerker, C.; Loeschinger, J.; Noda, M.; Bonhoeffer, F.  
Cell 82, 359-370, 1995  
A:Title: In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal pr  
A:Reference number: A57084; MUID:95360980

```

RESULT 7
138850
LEKK-4 - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
C:Accession: I38850
R:Kozlowsky, C.J.; Maraskovsky, E.; McGrew, J.T.; VandenBos, T.; Teepe, M.;

```

35	QY	YVWSSN-OHLRREGYTVQVNVNDYLDIYCPHYNSSGVGPAGPGGGAGQYVLYMYSR	93
		: : : : :     : : : : :     : : : : :	
33	Db	IYWNSSKFLPCQGLVLYPQIGDKDLIICPKVDSKTVGQ-----YKYKVMVDK	83
		: : : : :     : : : : :     : : : : :	
94	QY	NGYRTCNASQGFKECNRP----HAPHSPKFESEKFQYSAFSLGYEFHAGHEYYYIST	149
		: : : : :     : : : : :     : : : : :     : : : : :	
84	Db	DOADRCTIKK-----NTPLLNCAKPOODIKFTKQGFSPNLWGLFQKKNKDYIIIST	137
		: : : : :     : : : : :     : : : : :     : : : : :	
150	QY	PTHNLH-----WKCLRMKVPCVCCASHSGEKVPVLPQFTMGPNVKINVL	195
		: : : : :     : : : : :     : : : : :     : : : : :	
138	Db	SNGSLEGLDNQGGVCQTRAMKIL-MKVQODASSAGSTRNKDPTRRPELEAGTNGRSPT	196
		: : : : :     : : : : :     : : : : :     : : : : :	
196	QY	EDFEENPOVPKLESI--SGTSPKREHLPLAVGIAFLMTEL	236
		: : : : :     : : : : :     : : : : :     : : : : :	
197	Db	SPFVFPNPGSSTDGNSAGHSNNILGSEVALFAGIASGCIIFI	239
		: : : : :     : : : : :     : : : : :     : : : : :	



R:Shao, H.; Lou, L.; Pandey, A.; Pasquale, E.B.; Dixit, V.M.

J. Biol. Chem. 269, 26606-26609, 1994

A:Title: cDNA cloning and characterization of a ligand for the Cek5 receptor protein-tyr

A:Reference number: A5062; MUID:95014510

A:Accession: A5062

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-89, 'T', '91-345 <SHA>

A:CROSS-references: GB:U12983; NID:g575928; PIDN:AAA53231.1; PID:g575929

C:Genetics:

A:Gene: EPLG2

Query Match 10.7%; Score 139.5; DB 2; Length 345;  
Best Local Similarity 24.2%; Pred. No. 2.4e-05;  
Matches 59; Conservative 27; Mismatches 87; Indels 71; Gaps 8;

QY 4 APLLLLLLVVPLPLLAQGGGALGNHAYVWNSN-OHLRRGGYTVQVNVNDYLDIY 62

Db 15 AMVLTICRLATPLAK-----NLEPVSWSSLNPKFLSGKGLVIYPKIGDKLDII 63

QY 63 CPHYNSGVGPGAGPGGGAEQYVLYMV-----SRNGYRTCNASQGFKRWEC 110

Db 64 CPRAEA-----GRPEYKLYLVRPEQAACSTVLDPNVLVTN----- 102

QY 111 NRPHAPSPKPFSEKFORYSAFSLGVEFHAGHEYYISTPTNHLH-----WKCLRM 161

Db 103 ----KPHQEIIRFTIKFOEFSPNYMGLEPKKYHDYITSTNSGLENGREGGVCRTRM 158

QY 162 KVFCVCA-----STSHSGEKVPPTLPQFTMGPNVKINVLDFEGENPOVPKLEK 210

Db 159 KIVKVGDPNAVTPPEQLTTSRPSKESDNTVKTATQAPG---RGSQSDSGKHETVNOQEK 216

QY 211 SISG 214

Db 217 SGPG 220

RESULT 13

158406

LERK-2 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-Nov-1999

A:Accession: F58406

R:Fletcher, F.A.; Carpenter, M.; Shilling, H.; Baum, P.; Ziegler, S.; Holln

Oncogene 9, 3241-3248, 1994

A:Title: LERK-2, a ligand for the receptor tyrosine kinase ELK, is evolutionarily cons

A:Reference number: 158406; MUID:95022634

A:Accession: 158406

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-345 <RES>

A:CROSS-references: EMBL:U07560; NID:g563118; PIDN:AAA53092.1; PID:g563119

C:Genetics:

A:Gene: Eplg2

Query Match

Best Local Similarity 10.2%; Score 132.5; DB 2; Length 345;

Matches 58; Conservative 27; Mismatches 88; Indels 71; Gaps 8;

QY 4 APLLLLLLVVPLPLLAQGGGALGNHAYVWNSN-OHLRRGGYTVQVNVNDYLDIY 62

Db 15 AMVLTICRLATPLAK-----NLEPVSWSSLNPKFLSGKGLVIYPKIGDKLDII 63

QY 63 CPHYNSGVGPGAGPGGGAEQYVLYMV-----SRNGYRTCNASQGFKRWEC 110

Db 64 CPRAEA-----GRPEYKLYLVRPEQAACSTVLDPNVLVTN----- 102

QY 111 NRPHAPSPKPFSEKFORYSAFSLGVEFHAGHEYYISTPTNHLH-----WKCLRM 161

Db 103 ----KPHQEIIRFTIKFOEFSPNYMGLEPKKYHDYITSTNSGLENGREGGVCRTRM 158

QY 162 KVFCVCA-----STSHSGEKVPPTLPQFTMGPNVKINVLDFEGENPOVPKLEK 210

Db 159 KIVKVGDPNAVTPPEQLTTSRPSKESDNTVKTATQAPG---RGSQSDSGKHETVNOQEK 216

QY 211 SISG 214

Db 217 SGPG 220

RESULT 14

T40265

hypothetical zinc-finger protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T40265

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z21868

A:Accession: T40265

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-356 <LN>

A:CROSS-references: EMBL:AL031854; PIDN:CAA21282.1; GSPDB:GN00067; SPDB:SPBC337.12

A:Experimental source: strain 972h-; cosmid c337

C:Genetics:

A:Gene: SPDB:SPBC337.12

A:Map position: 2

Query Match

Best Local Similarity 7.7%; Score 100; DB 2; Length 356;

Matches 42; Conservative 31; Mismatches 64; Indels 84; Gaps 8;

QY 35 VYNNSSNQHLRRGGYTVQVNVNDYLDIYCPHYNSGVGPGAGPGGGAEQYVLYMVSR- 93

Db 161 LYNNKSYLLKKRFLKEVG-NSPSAVYCRYNANGI-----CGKGAACRFVHEPTRK 212

QY 94 -----NGYRTCN-----ASQGFKRWCNRPAPHSPKIFSE-----K 125

Db 213 TICPKFLNGRCNAEDCNLSHEDPRIPACRYFLLLGCKNNPCRYVHIHSENAPICFE 272

QY 126 FQYSAFSLGVEFHAGHEYYISTPTNHLH-WKCLRMKVFVCCASTSHSGEKVPPTLPQF 184

Db 273 FAKYFGCELG-----TSCKNQHLQCTDYAMFGSC----- 302

QY 185 TMGPNVKINVLDFEGENPOVPKLEKSISCTSPKREHLPLA 225

Db 303 -----NNFOCSLYHGAVSADVPQEAPIS 327

RESULT 15

GNV27

genome polyprotein - human poliovirus 3 (strain 23127)

N:Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4;

A-directed RNA polymerase (EC 2.7.7.48)

C:Species: human poliovirus 3

C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 16-Jul-1999

C:Accession: A27245

R:Hughes, P.J.; Evans, D.M.A.; Minor, P.D.; Schild, G.C.; Almond, J.W.; Stanway, G.

J. Gen. Virol. 67, 2093-2102, 1986

A:Title: The nucleotide sequence of a type 3 poliovirus isolated during a recent out

A:Reference number: A27245; MUID:87010550

A:Accession: A27245

A:Molecule type: genomic RNA

A:Residues: 1-2206 <HUG>

A:CROSS-references: GB:X04468; NID:g61112; PIDN:CAA28155.1; PID:g61113

C:Superfamily: poliovirus genome polyprotein

C:Keywords: coat protein; core protein; genome-linked protein; nucleotidyltransferas

F:1-69/Product: coat protein VP4 #status predicted <VP4>

F:70-340/Product: coat protein VP2 #status predicted <VP2>

F:341-578/Product: coat protein VP3 #status predicted <VP3>

F:579-878/Product: coat protein VP1 #status predicted <VP1>

F:879-1027/Product: core protein P2-3b #status predicted <P3b>



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:53:40 ; Search time 24.32 seconds  
(without alignments)  
239.033 Million cell updates/sec

Title: US-09-733-756-2

Perfect score: 1301

Sequence: 1 MAAAPLLLLLLVPVPLPL.....REHLPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	100.0	238	1	US-08-240-124-2
2	1301	100.0	238	1	US-08-453-943-2
3	1301	100.0	238	2	US-09-057-121-2
4	1301	100.0	238	4	US-09-358-734-2
5	1261	96.9	234	1	US-08-299-567-5
6	437	33.6	209	1	US-08-455-001-2
7	437	33.6	209	4	US-08-308-814-2
8	437	33.6	209	5	PCT-US95-11869-2
9	437	33.6	213	1	US-09-609-324A-10
10	437	33.6	213	2	US-08-920-440B-10
11	437	33.6	213	4	US-09-173-492-10
12	437	33.6	213	4	US-09-173-133-10
13	421	32.4	200	1	US-08-455-001-4
14	421	32.4	200	5	PCT-US95-11869-4
15	415.5	31.9	228	1	US-08-442-248-4
16	415.5	31.9	228	1	US-08-440-815-4
17	415.5	31.9	228	3	US-08-379-802-2
18	415.5	31.9	228	3	US-09-048-129-2
19	415.5	31.9	228	4	US-09-048-079-2
20	415.5	31.9	228	4	US-08-486-449-4
21	415.5	31.9	228	5	PCT-US95-15781-5
22	403	31.0	184	1	US-09-609-324A-2
23	403	31.0	184	2	US-08-920-440B-2
24	403	31.0	184	4	US-09-173-492-2
25	403	31.0	184	4	US-09-173-133-2
26	403	31.0	184	4	US-09-165-533-2
27	403	31.0	184	5	PCT-US95-12779-2

Sequence 2, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 4, Appli  
Sequence 1, Appli  
Sequence 12, Appli  
Sequence 1, Appli  
Sequence 12, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 8, Appli  
Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-08-240-124-2  
; Sequence 2, Application US/08240124  
; Patent No. 5516658  
; GENERAL INFORMATION:  
; APPLICANT: BECKMANN, M. P.  
; APPLICANT: CERRETI, DOUGLAS P.  
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
; TITLE OF INVENTION: RECEPTOR HEX  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMUNEX CORPORATION  
; STREET: 51 UNIVERSITY STREET  
; CITY: SEATTLE  
; STATE: WASHINGTON  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Apple System 7.1  
; SOFTWARE: Microsoft Word for Apple, Version 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/240,124  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: US 08/161,132  
; APPLICATION NUMBER: US 08/161,132  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/114,426  
; FILING DATE: 30-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/109,745  
; FILING DATE: 20-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SEESE, KATHRYN A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2814-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 238 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-240-124-2

```
; MOLECULE TYPE: protein
US-08-453-943-2
```

Query Match 100.0%; Score 1301; DB 1; Length 238;  
Best Local Similarity 100.0%; Pred. No. 2.7e-128;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLPLLLLVVPVLPDLAAGPGGALGNRRHAVYMSSNOHLRREGYTVOVNNDYLD 60  
| | | | |  
DB 1 MAAPLPLLLLVVPVLPDLAAGPGGALGNRRHAVYMSSNOHLRREGYTVOVNNDYLD 60  
| | | | |  
QY 61 IYCPHYNSSGVGGACGPVGGAEOVLVMVRNGRYTCNASQGFKRWECNRPAPHSPHPI 120  
| | | | |  
DB 61 IYCPHYNSSGVGGACGPVGGAEOVLVMVRNGRYTCNASQGFKRWECNRPAPHSPHPI 120  
| | | | |  
QY 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNLHWKLRMKVVFCCASTSHSGEKPVT 180  
| | | | |  
DB 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNLHWKLRMKVVFCCASTSHSGEKPVT 180  
| | | | |  
QY 181 LPQFTMGPNVKINLVDFEGENPQPVKLEKSISGTSPKRHLPLAVGIATFLMTFLAS 238  
| | | | |  
DB 181 LPQFTMGPNVKINLVDFEGENPQPVKLEKSISGTSPKRHLPLAVGIATFLMTFLAS 238  
| | | | |

RESULT 3  
US-09-057-121-2  
Sequence 2, Application US/09057121  
Patent No. 5969110  
GENERAL INFORMATION:  
APPLICANT: BECKMANN, M. P.  
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
TITLE OF INVENTION: RECEPTOR HEK  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMUNEX CORPORATION  
STREET: 51 UNIVERSITY STREET  
CITY: SEATTLE  
STATE: WASHINGTON  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Apple Macintosh  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/057,121  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/240,124  
FILING DATE:  
APPLICATION NUMBER: US 08/161,132  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/114,426  
FILING DATE: 30-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/109,745  
FILING DATE: 20-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SEESE, KATHRYN A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2814-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEX: (206) 233-0644  
TELEX: 756822  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 238 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-057-121-2

Query Match 100.0%; Score 1301; DB 2; Length 238;  
Best Local Similarity 100.0%; Pred. No. 2.7e-128;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAAPLLLLLLVPPVPLPPLLAQGGGALGNRHAYVWSSNOHLRREGYTVQVNVNDYLD 60  
Db 1 MAAAPLLLLLLVPPVPLPPLLAQGGGALGNRHAYVWSSNOHLRREGYTVQVNVNDYLD 60  
QY 61 IYCPHYNSSGVGPGAGPGGGAEOYVLYMVSRNGYRTCTNASQGFKRWECNRPAPHSP 120  
Db 61 IYCPHYNSSGVGPGAGPGGGAEOYVLYMVSRNGYRTCTNASQGFKRWECNRPAPHSP 120  
QY 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVYVCCASTSHSGEKPVP 180  
Db 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVYVCCASTSHSGEKPVP 180  
QY 181 LPQFTMGPNVKINVLDEFEENPQVPKLEKSIKSTSPKREHLPLAVGIAFFLMTFLAS 238  
Db 181 LPQFTMGPNVKINVLDEFEENPQVPKLEKSIKSTSPKREHLPLAVGIAFFLMTFLAS 238

RESULT 4  
US-09-358-734-2  
; Sequence 2, Application US/09358734  
; Patent No. 6274117  
; GENERAL INFORMATION:  
; APPLICANT: BECKMANN, M. P.  
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
; TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMUNEX CORPORATION  
; STREET: 51 UNIVERSITY STREET  
; CITY: SEATTLE  
; STATE: WASHINGTON  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Apple System 7.1  
; SOFTWARE: Microsoft Word for Apple, Version 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/358,734  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/240,124  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/114,426  
; FILING DATE: 30-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/109,745  
; FILING DATE: 20-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SEESE, KATHRYN A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2814-C  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 238 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-358-734-2

Query Match 100.0%; Score 1301; DB 4; Length 238;  
Best Local Similarity 100.0%; Pred. No. 2.7e-128;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAAPLLLLLLVPPVPLPPLLAQGGGALGNRHAYVWSSNOHLRREGYTVQVNVNDYLD 60  
Db 1 MAAAPLLLLLLVPPVPLPPLLAQGGGALGNRHAYVWSSNOHLRREGYTVQVNVNDYLD 60  
QY 61 IYCPHYNSSGVGPGAGPGGGAEOYVLYMVSRNGYRTCTNASQGFKRWECNRPAPHSP 120  
Db 61 IYCPHYNSSGVGPGAGPGGGAEOYVLYMVSRNGYRTCTNASQGFKRWECNRPAPHSP 120  
QY 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVYVCCASTSHSGEKPVP 180  
Db 121 KFSEKFORYSFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVYVCCASTSHSGEKPVP 180  
QY 181 LPQFTMGPNVKINVLDEFEENPQVPKLEKSIKSTSPKREHLPLAVGIAFFLMTFLAS 238  
Db 181 LPQFTMGPNVKINVLDEFEENPQVPKLEKSIKSTSPKREHLPLAVGIAFFLMTFLAS 238

RESULT 5  
US-08-299-567-5  
; Sequence 5, Application US/08299567  
; Patent No. 5747033  
; GENERAL INFORMATION:  
; APPLICANT: Davis, et al.  
; TITLE OF INVENTION: METHOD OF ENHANCING THE BIOLOGICAL  
; TITLE OF INVENTION: ACTIVITY OF EPH FAMILY LIGANDS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-6707  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/299,567  
; FILING DATE: 01-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kempler, Gail M.  
; REGISTRATION NUMBER: 32,143  
; REFERENCE/DOCKET NUMBER: REG 290  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 914-345-7400  
; TELEFAX: 914-345-7721  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 234 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-299-567-5  
Query Match 96.9%; Score 1261; DB 1; Length 234;  
Best Local Similarity 97.9%; Pred. No. 4e-124;  
Matches 233; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 MAAAPLLLLLLVPVPLPALLAOGPGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLD 60  
Db 1 MAAAPLLLLLLVPVPLPALLAOGPGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLD 60  
QY 61 IYCPHYNSSGVGPGAGPGGGAEOYVLYWVSRNGYRTCNASQGFKRWCNRPHPHPSPI 120  
Db 61 IYCPHYNSS----GAGPGGGGAEOYVLYWVSRNGYRTCNASQGFKRWCNRPHPHPSPI 116  
QY 121 KFSEKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVT 180  
Db 117 KFSEKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKPVT 176  
QY 181 LPOFTMGPNVKNVLEDFEGENPQVPKLEKSISGTSKREHPLAVGIAFFLMTFLAS 238  
Db 177 LPOFTMGPNVKNVLEDFEGENPQVPKLEKSISGTSKREHPLAVGIAFFLMTFLAS 234

## RESULT 6

US-08-455-001-2  
; Sequence 2, Application US/08455001  
; Patent No. 5795734  
; GENERAL INFORMATION:  
; APPLICANT: Flanagan, John G.  
; APPLICANT: Cheng, Hwai-Jong  
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related  
; TITLE OF INVENTION: Thereeto  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/455.001  
; FILING DATE: 31 MAY 1995  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-011CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 209 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-455-001-2

Query Match 33.6%; Score 437; DB 1; Length 209;  
Best Local Similarity 50.5%; Pred. No. 5.4e-38;  
Matches 96; Conservative 25; Mismatches 47; Indels 22; Gaps 7;  
QY 2 AAAPLLLLLLVPVPLPALLA--OGPGALGNRHAVYWNSSNOHLRRE-----GYTVQV 53  
Db 4 AQRPLPLLLL-----LLPLARNEDPARANADRYAVYWNRSNPRFQVSAVGDGGGYTVEV 59  
QY 54 NVNDYLDIYCPHYNSSGVGPGAGPGGGAEOYVLYWVSRNGYRTCNASQ--GFKRWCN 112  
Db 60 SINDYLDIYCPHY-----GAPLPPAERMERYILYVWNGEGHASCDCRQGRGFKRWCN 112  
QY 113 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170  
Db 117 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170  
QY 177 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170  
Db 177 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170

Db 113 PAAPGGPLKFSEKFQLTFTPSLGFEPGHEYYIISATPPNLVDRPCLRLKVVVRPTNET 172  
QY 171 SHSGEKPVT 180  
Db 173 LYEAPPEIPT 182  
RESULT 7  
US-08-308-814-2  
; Sequence 2, Application US/08308814  
; Patent No. 6268476  
; GENERAL INFORMATION:  
; APPLICANT: Flanagan, John G.  
; APPLICANT: Cheng, Hwai-Jong  
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related  
; TITLE OF INVENTION: Thereeto  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (txt)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/308.814  
; FILING DATE: 19-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 209 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-308-814-2

Query Match 33.6%; Score 437; DB 4; Length 209;  
Best Local Similarity 50.5%; Pred. No. 5.4e-38;  
Matches 96; Conservative 25; Mismatches 47; Indels 22; Gaps 7;  
QY 2 AAAPLLLLLLVPVPLPALLA--OGPGALGNRHAVYWNSSNOHLRRE-----GYTVQV 53  
Db 4 AQRPLPLLLL-----LLPLARNEDPARANADRYAVYWNRSNPRFQVSAVGDGGGYTVEV 59  
QY 54 NVNDYLDIYCPHYNSSGVGPGAGPGGGAEOYVLYWVSRNGYRTCNASQ--GFKRWCN 112  
Db 60 SINDYLDIYCPHY-----GAPLPPAERMERYILYVWNGEGHASCDCRQGRGFKRWCN 112  
QY 113 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170  
Db 113 PHAPSPITKFSKQRYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCAS-T 170  
QY 171 SHSGEKPVT 180  
Db 173 LYEAPPEIPT 182  
RESULT 8  
PCR-US95-11869-2  
; Sequence 2, Application PC/TUS9511869

```

RESULT          9
US-09-609-324A-10
; Sequence 10, Application US/09609324A
; Patent No. RE37582
; GENERAL INFORMATION:
; APPLICANT: CERRETTI, Douglas P.
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6
; FILE REFERENCE: A7772
; CURRENT APPLICATION NUMBER: US/09/609,324A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/920,440
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 08/538,709
; PRIOR FILING DATE: 1995-10-03
; PRIOR APPLICATION NUMBER: 08/318,393
; PRIOR FILING DATE: 1994-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: LERK-6
US-09-609-324A-10

```

		Query Match	33.6%;	Score 437;	DB 2;	Length 213;	
		Best Local Similarity	50.5%;	Pred. No. 5.6e-38;			
		Matches	96;	Conservative	24;	Mismatches	52; Indels 18; Gaps 7;
QY	2	AAAPLL-LULLLPVPPLLPL-AQGPGGALGNRHAVYNSSNOHLRR-----EGYTVGV	53				
		:   :   :		:	: :		
Db	4	AQRPLP LULLLPLPPFPARFADRAANSDRIVATVWNRSPFRFIAGAGDCGGGYTVEV	63				
QY	54	NVNDYLDIYPHYNSSGVGPFGPGGGAEQVLVYMVSNGRYRTCTNASO-GFKRWECNR	112				
		: :                     :		:	:		

		Query Match	33.6%;	Score 437;	DB 2;	Length 213;	
		Best Local Similarity	50.5%;	Pred. No. 5.6e-38;			
		Matches	96;	Conservative	24;	Mismatches	52; Indels 18; Gaps 7;
QY	2	AAAPLL-LULLLPVPPLLPL-AQGPGGALGNRHAVYNSSNOHLRR-----EGYTVOV	53				
		:   :	: : : :				
Db	4	AQRPLP LULLLPLPPPFARADRAANSDRIVATVWNRSPFRFIAGAGDCGGGYTVEV	63				
QY	54	NVNDYLDIYPHYNSSGVGPFGPGGGAEQVLVYMVSNGRYRTCTNASO-GFKRWECNR	112				
		: : : :					

Db 64 SINDYLDIYCPHY-----GAPLPAERMEHYVLMVNGEGHASCDSRQGRQGRKRWECNR 116  
QY 113 PHAPHSPIKSEKQRYSAFSLGYEFHAGHEYYIIS-TPTHNLHWKCLRMKVFVCCAS-T 170  
Db 117 PAAPGGPLAKSEKQFLTFPSLGFEPGHEYYIISATPPNAVDRPCLRLKVVYRPTNET 176  
QY 171 SHSGEKPVPT 180  
Db 177 LYEAPPIFT 186  
RESULT 11  
US-09-173-492-10  
; Sequence 10, Application US/09173492  
; Patent No. 6194172  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,492  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 213 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-173-492-10

Query Match 33.6%; Score 437; DB 4; Length 213;  
Best Local Similarity 50.5%; Pred. No. 5.6e-38;  
Matches 96; Conservative 24; Mismatches 52; Indels 18; Gaps 7;  
QY 2 AAAPLL-LLLLVPVPLPLL-AQPGGALGNRHAVYVWNSNQHLRR-----EGYTVQV 53  
Db 4 AORPLLPLLLLLPLPPPPFAAEADAARANSRDYAVYVWNSNPRFHAGAGDDGGYTV 63  
QY 54 NVNDYLDIYCPHYNSGVGPGAGPGGAGQYVLYVWVSRNGYRTCNASQ-GFKRWECNR 112  
Db 64 SINDYLDIYCPHY-----GAPLPAERMEHYVLMVNGEGHASCDSRQGRKRWECNR 116  
QY 113 PHAPHSPIKSEKQRYSAFSLGYEFHAGHEYYIIS-TPTHNLHWKCLRMKVFVCCAS-T 170  
Db 117 PAAPGGPLKSEKQFLTFPSLGFEPGHEYYIISATPPNAVDRPCLRLKVVYRPTNET 176  
QY 171 SHSGEKPVPT 180  
Db 177 LYEAPPIFT 186

RESULT 13  
US-08-455-001-4  
; Sequence 4, Application US/08455001  
; Patent No. 5795734  
; GENERAL INFORMATION:  
; APPLICANT: Flanagan, John G.  
; APPLICANT: Cheng, Hwai-Jong

Query Match 33.6%; Score 437; DB 4; Length 213;  
Best Local Similarity 50.5%; Pred. No. 5.6e-38;  
Matches 96; Conservative 24; Mismatches 52; Indels 18; Gaps 7;

QY 2 AAAPLL-LLLLVPVPLPLL-AQPGGALGNRHAVYVWNSNQHLRR-----EGYTVQV 53  
Db 4 AORPLLPLLLLLPLPPPPFAAEADAARANSRDYAVYVWNSNPRFHAGAGDDGGYTV 63  
QY 54 NVNDYLDIYCPHYNSGVGPGAGPGGAGQYVLYVWVSRNGYRTCNASQ-GFKRWECNR 112  
Db 64 SINDYLDIYCPHY-----GAPLPAERMEHYVLYVWVNGEGHASCDSRQGRKRWECNR 116  
QY 113 PHAPHSPIKSEKQRYSAFSLGYEFHAGHEYYIIS-TPTHNLHWKCLRMKVFVCCAS-T 170  
Db 117 PAAPGGPLKSEKQFLTFPSLGFEPGHEYYIISATPPNAVDRPCLRLKVVYRPTNET 176  
QY 171 SHSGEKPVPT 180  
Db 177 LYEAPPIFT 186

5

**ATTORNEY/AGENT INFORMATION:**

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-442-248-4

Query Match      31.9%; Score 415.5; DB 1; Length 228;
Best Local Similarity 39.5%; Pred. No. 1.1e-35;
Matches 98; Conservative 37; Mismatches 66; Indels 47; Gaps 11;

QY 8 LLLLLVPVLLPLLAQPGG-ALGNRHAVYWNSSNQHLRREGYTVOVWVNDYLDIYCPHY 66
Db 6 MLTLVFLVLWMCVFSQDPGSKAVADRYAVYWNSSNPRFQRGDYHIDVCINDYLDVFCPHY 65

QY 67 NSSGVGPGAGPGPGGAEQYVLYMYSRNGYRTC-NASQGFKRWECNRPHAPHSPIKFSEK 125
Db 66 EDS-----VPEDKTERYVLYMVNFDGYSACDHTSKGFKRWEKNRPHSPNGPLKFSEK 117

QY 126 FQYSAFSLGYEFHAGHEYYIISTP-THNLHWKCLRMKVYV----CCASTSHSGEKPVP 180
Db 118 FQLTFPFSIGFEFRPGREYFISSAIPDNGRRSCLKLKVFRPTNSCMKTIQVHDR---- 173

QY 181 LPOFTMGPNVKINLEDFE-----GEN-POVPKLEKSISGTSKREHLPLAVGI 228
Db 174 --VFDVNDKVE-NSLEPADDTVHESAEPSSRGENAAQTPI-----PSR-----LLAI 217

QY 229 AFFLMTFL 236
Db 218 LLEFLAML 225
```

Search completed: July 13, 2002, 10:45:32  
Job time: 3112 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2002, 09:49:45 ; Search time 57.73 seconds  
(without alignments)  
457.917 Million cell updates/sec

Title: US-09-733-756-2

Perfect score: 1301

Sequence: 1 MAAPLLLLLLVPPVLLPL.....REHPLAVGIAFFLMTFLAS 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1301	100.0	238	16 AAR71481	Human hek-L protei
2	1266	97.3	234	16 AAR82605	Eph transmembrane
3	437	33.6	209	17 AAR94766	Mouse Eph receptor
4	437	33.6	209	19 AAR71006	Amino acid sequenc
5	437	33.6	213	20 AAY06822	Human LERK-6 polyp
6	426.5	32.8	335	22 AAG27837	Novel human diagno
7	421	32.4	200	17 AAR94767	Chicken Eph recept
8	421	32.4	200	19 AAR71007	Amino acid sequenc
9	415.5	31.9	228	17 AAR97854	Human AL-1, a liga
10	415.5	31.9	228	17 AAW02586	Lerk-7 protein. H
11	415.5	31.9	228	17 AAW00035	HEK4 binding prote

12	403	31.0	184	17 AAW02587	Lerk-6 protein. M
13	403	31.0	184	17 AAR91283	LERK-6 protein. M
14	403	31.0	184	20 AAY06820	Murine LERK-6 poly
15	390	30.0	205	13 AAR23895	B61 protein sequen
16	390	30.0	205	15 AAR53634	EBP. Homo sapiens
17	390	30.0	205	16 AAR82604	Eph transmembrane
18	390	30.0	205	18 AAW04632	eck receptor bindi
19	386	29.7	205	22 AAS05090	Human PRO202 prote
20	386	29.7	218	21 AAS54058	Human pancreatic c
21	384	29.5	180	18 AAW18082	Truncated eck rece
22	382	29.4	171	18 AAW18081	Truncated eck rece
23	380	29.2	167	18 AAW18080	Truncated eck rece
24	379.5	29.2	160	18 AAW18079	Truncated eck rece
25	377.5	29.0	151	18 AAW18078	Truncated eck rece
26	377.5	29.0	204	22 AAU12315	Human PRO202 poly
27	377.5	29.0	204	22 AAB50974	Human PRO202 prote
28	364.5	28.0	179	17 AAR94768	EPH receptor ligand
29	364.5	28.0	179	19 AAR71008	Generic sequence f
30	364.5	28.0	201	16 AAR71482	Human hek-L protei
31	334	25.7	104	17 AAR91284	LERK-6 exon polype
32	334	25.7	104	21 AAY06821	Amino acid sequenc
33	295	22.7	58	21 AAG03418	Human secreted pro
34	180	13.8	340	18 AAW17081	EPH family ligand
35	180	13.8	340	19 AAW46615	Human transmembran
36	179	13.8	340	18 AAW33699	AL-2-short (AL-2s)
37	179	13.8	340	18 AAW31544	Human cytokine Ler
38	179	13.8	340	18 AAW10637	NLERK2 ligand for
39	179	13.8	455	18 AAW33698	AL-2-long (AL-2l)
40	177	13.6	658	21 AAY96782	Ephrin-B2-Ephrin-B
41	171.5	13.2	308	17 AAW06334	Ligand #2 for rece
42	171.5	13.2	308	17 AAW94656	Ligand for recepto
43	171.5	13.2	333	17 AAW06337	Full length ligand
44	171.5	13.2	333	17 AAR94655	Ligand for recepto
45	171.5	13.2	333	17 AAR92743	Human hepatoma tra

#### ALIGNMENTS

RESULT 1  
AAR71481  
ID AAR71481 standard; Protein; 238 AA.  
XX  
AC AAR71481;  
XX  
DT 03-OCT-1995 (first entry)  
XX  
DE Human hek-L protein.  
XX  
KW Ligand; cell surface; tyrosine kinase receptor; tumorigenesis;  
KW Immunogen.  
XX  
OS Homo sapiens.  
FH Key  
FT Peptide  
FT Peptide  
FT Protein  
FT Protein  
PN MO9506065-A  
XX  
PD 02-MAR-1995.  
XX  
PF 17-AUG-1994; 94WO-US09282.  
XX  
PR 20-AUG-1993; 93US-0109745.  
PR 30-AUG-1993; 93US-0114426.  
PR 03-DEC-1993; 93US-0161132.  
PR 09-MAY-1994; 94US-0240124.  
XX  
PA (IMMV ) IMMUNEX CORP.  
XX Beckmann MP, Cerretti DP;  
PI

XX DR WPI; 1995-106811/14.  
 XX N-PSDB; AAQ85887.  
 XX PT New isolated DNA encoding hek-L protein or its fusion products -  
 XX PT useful as assay reagent or for carrying therapeutic and  
 XX PT diagnostic compounds to leukaemia cells.  
 XX PS Claim 21; Page 36; 45pp; English.  
 XX CC The sequence is that of a novel protein designated hek-L, a protein  
 XX CC that can bind hek (a cell surface receptor tyrosine kinase). Hek-L  
 XX CC is the first known ligand for hek and can be used to study cellular  
 XX CC processes regulated by hek (which may be involved in tumorigenesis).  
 XX CC It is also an immunogen for antibody production, as a reagent for  
 XX CC detecting hek or hek-L in in vitro assays, to determine binding of  
 XX CC hek proteins, to purify hek proteins, and to carry diagnostic or  
 XX CC cytotoxic agents to particular leukaemia cells that express the hek  
 XX CC antigen. Hek-L also binds the elk tyrosine kinase receptors.  
 XX CC See also AAR71482.  
 XX SQ Sequence 238 AA;  
 CC  
 Query Match 100.0%; Score 1301; DB 16; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 8,1e-124;  
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAAPLPLLLLLVLPVPLPLLAQGGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLD 60  
 DB 1 maapllllllllvvpvplllaqpggaglnrhavynssnqhlrregyvtqvnvndyld 60  
 QY 61 IYCPHYNSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCNASQGFKRWECHRHAPHSP 120  
 DB 61 IYCPHYNSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCNASQGFKRWECHRHAPHSP 120  
 QY 121 KFSEKFORYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKVP 180  
 DB 121 kfsekforysafslgyefhagheyiystptnhlhwkclrmkvfvccastshsgekvp 180  
 QY 181 LPQFTMGPNVKINVLDEFEENPQVPKLEISGTSFKREHLPLAVGIAFFLMTFLAS 238  
 DB 181 lpqftmgpnvkinvledefeenpqvpklesigtsfkrehlplavgi afflmtflas 238  
 RESULT 2  
 AAR82605  
 ID AAR82605 standard; Protein: 234 AA.  
 AC AAR82605;  
 XX 16-MAY-1996 (first entry)  
 DE Eph transmembrane tyrosine kinase family ligand, Efl-2.  
 KW Efl-2; EHK1-L; Eph transmembrane tyrosine kinase family ligand;  
 KW neurological disorder; identification; diagnosis.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Peptide 1..30  
 FT Region /label= signal\_peptide  
 FT 218..235  
 FT /note= "carboxy terminal hydrophobic  
 FT GPI-recognition tail"  
 FT Misc-difference 158  
 FT /note= "residue borders main conserved regions"  
 XX  
 PN W09527060-A2.  
 XX 12-OCT-1995.  
 PD  
 XX

PF 04-APR-1995; 95WO-US04208.  
 XX  
 XX 21-OCT-1994; 94US-0327423.  
 PR 04-APR-1994; 94US-0222075.  
 PR 12-APR-1994; 94US-0229402.  
 PR 01-SEP-1994; 94US-0299567.  
 XX  
 PA (REGE-) REGENERON PHARM INC.  
 XX  
 XX Aldrich TH, Davis S, Gale N, Goldfarb M, Maisonpierre PC;  
 PI Yancopoulos GD;  
 XX  
 DR WPI; 1995-358635/46.  
 DR N-PSDB; AAT03883.  
 XX  
 XX Ligands which bind Eph family receptors - used in the diagnosis of  
 PT neurological disorders  
 XX  
 PS Disclosure; Fig 2; 58pp; English.  
 XX  
 CC Efl-2 (also known as EHK-1L) is an Eph transmembrane tyrosine kinase  
 CC family ligand. It has homology with B61 (Efl-1) (see AAR82604). Efl-2  
 CC ends in a C-terminal hydrophobic sequence that appears to be a  
 CC recognition sequence allowing it to be GPI-linked and thus lacking in  
 CC an intracellular domain. Efl-2 is useful for identifying other ligands  
 CC for EHK-1, -2, -3, Eck and Elk receptors. The ligands are useful in  
 CC promoting a differential function and/or influencing the phenotype,  
 CC such as growth and/or proliferation, of receptor bearing cells. They  
 CC may be used in the diagnosis, and treatment of neurological disorders.  
 XX  
 XX Sequence 234 AA;  
 SQ  
 Query Match 97.3%; Score 1266; DB 16; Length 234;  
 Best Local Similarity 98.3%; Pred. No. 2.8e-120;  
 Matches 234; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
 QY 1 MAAPLPLLLLLVLPVPLPLLAQGGGALGNRHAVYWNSSNOHLRREGYTVQVNVNDYLD 60  
 DB 1 maapllllllllvvpvplllaqpggaglnrhavynssnqhlrregyvtqvnvndyld 60  
 QY 61 IYCPHYNSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCNASQGFKRWECHRHAPHSP 120  
 DB 61 IYCPHYNSSGVGPGAGPGGGAGQYVLYMYSRNGYRTCNASQGFKRWECHRHAPHSP 120  
 QY 121 KFSEKFORYSAFSLGYEFHAGHEYIYSTPTNHLHWKCLRMKVFVCCASTSHSGEKVP 180  
 DB 117 kfsekforysafslgyefhagheyiystptnhlhwkclrmkvfvccastshsgekvp 176  
 QY 181 LPQFTMGPNVKINVLDEFEENPQVPKLEISGTSFKREHLPLAVGIAFFLMTFLAS 238  
 DB 177 lpqftmgpnvkinvledefeenpqvpklesigtsfkrehlplavgi afflmtflas 234  
 RESULT 3  
 AAR94766  
 ID AAR94766 standard; Protein: 209 AA.  
 XX  
 XX AAR94766;  
 XX  
 XX 02-JUL-1996 (first entry)  
 DT Mouse Eph receptor ligand Efl-1.  
 DE  
 DE Mouse Eph receptor ligand Efl-1.  
 KW Efl-1; Eph receptor ligand; dementia; tachycardia; therapy;  
 KW diagnosis; transgenic animal.  
 XX  
 XX Mus musculus.  
 OS  
 XX Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= sig\_peptide  
 FT Protein 21..209



QY 2 AAAPLLLLLVPPVPLLLA--QGGGALGNRHAVYWNSSNOHLRR-----GYTVQV 53  
 Db 4 aqrpllp1lll1---llplarnedparanadyaywnsrprfvsavgdgggytvev 59  
 QY 54 NVNDYLDIYCPHYNSGVGPGGGAQYVLYVWNRNGYRTCNASQ-GFKRWECNR 112  
 Db 60 sindyldiycphy-----gaplpaaermeryllymngvgehascdhrqgrfkrwecnr 112  
 QY 113 PHAPSPKIFSEKFORYSAFSLGVEFHAGHEYIIS-TPTHNLHWKCLRMKVFVCCAS-T 170  
 Db 113 paapggplkfsekqlftpslgsfgrphgheyiyisatpnlvdrcilrkvyvrtnet 172  
 QY 171 SHSGEKPVPPT 180  
 Db 173 lyeapepift 182

## RESULT 5

AA06822  
 ID AAY06822 standard; Protein; 213 AA.

XX AC AAY06822;

XX DT 24-JUN-1999 (first entry)

XX DE Human LERK-6 polypeptide.

XX KW LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;  
 cell proliferation; neural growth; neural tissue; neurological disease;  
 neurodegenerative; excitotoxicity.

XX OS Homo sapiens.

XX PN WO9910495-A1.

XX PD 04-MAR-1999.

XX PF 27-AUG-1998; 98WO-US17772.

XX PR 29-AUG-1997; 97US-0920440.

XX PA (IMMV ) IMMUNEX CORP.

XX PI Cerretti DP;

XX DR WPI; 1999-243567/20.

XX DR N-PSDB; AAX32767.

XX PT New cytokine designated LERK-6

XX PS Claim 6; Page 42; 46pp; English.

XX CC The invention relates murine and human LERK-6 polypeptides that bind to  
 hek/elk receptors. Host cells transfected or transformed with vectors  
 comprising the LERK-6 nucleic acid sequences are used for the recombinant  
 production of the proteins. LERK-6 polypeptides may be useful in the  
 enhancement, stimulation, proliferation or growth of cells expressing the  
 hek or elk receptor. The ligand and receptor complex may be involved in  
 neural growth, development and/or maintenance. LERK-6 can be used for  
 treating disorders of neural tissue such as injury or neurological  
 diseases, either chronic or acute. LERK-6 may be employed in treating  
 neurodegenerative conditions where there is neural death, excitotoxicity.  
 In addition, they may be administered to a mammal to exert a trophic  
 effect on neural tissue. They can also be used as reagents for those  
 conducting quality assurance studies e.g. to monitor shelf life and  
 stability of elk protein under different conditions. The polypeptides can  
 also be used as carriers for delivering agents attached to cells bearing  
 the elk or hek cell surface receptor. The present sequence represents a  
 human LERK-6 polypeptide.

XX Sequence 213 AA;

XX SQ

Query Match 33.6%; Score 437; DB 20; Length 213;  
 Best Local Similarity 50.5%; Pred. No. 3.7e-36;  
 Matches 96; Conservative 24; Mismatches 52; Indels 18; Gaps 7;

QY 2 AAAPLL-LLLLLVPPVPLLLA--AQQPGGALGNRHAVYWNSSNOHLRR-----EGYTVQV 53  
 Db 4 aqrpllp1lll1lllpppfaraedaarandsryaywnsrprfhagagddgggytvev 63  
 QY 54 NVNDYLDIYCPHYNSGVGPGGGAQYVLYVWNRNGYRTCNASQ-GFKRWECNR 112  
 Db 64 sindyldiycphy-----gaplpaaermehylymngvgehascdhrqgrfkrwecnr 116  
 QY 113 PHAPSPKIFSEKFORYSAFSLGVEFHAGHEYIIS-TPTHNLHWKCLRMKVFVCCAS-T 170  
 Db 117 paapggplkfsekqlftpslgsfgrphgheyiyisatpnlvdrcilrkvyvrtnet 176  
 QY 171 SHSGEKPVPPT 180  
 Db 177 lyeapepift 186

## RESULT 6

ABG27837  
 ID ABG27837 standard; Protein; 335 AA.

XX AC ABG27837;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #27828.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS92024.

XX CC New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity

XX PS Claim 20; SEQ ID NO 58196; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving  
 (II). (II) is useful for generating antibodies against it, detecting or  
 quantitating a polypeptide in tissue, as molecular weight markers and as  
 a food supplement. (II) and its binding partners are useful in medical  
 imaging of sites expressing (II). (I) and (II) are useful for treating  
 disorders involving aberrant protein expression or biological activity.  
 The polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations

XX



Qy	8	LLLLLVVPVPLPLLAQPGG-ALGNRHAVYWNSSNOHLRREGTIVQVNVNDYLDIYCPHY	66
Db	6	mtlvtvlvmcvfsqdgskavadrYavvwnsnprfgrdyhldvcindyldvfcphy	65
Qy	67	NSSGVGPGAGPGCGGAEQVLYVMVSRNGYRTC-NASOGFKRWECHRPAPHSPKIFSEK	125
Db	66	eds-----vpedkterylvymvfdgysacdhstskgfkwechrphspnpgplkfsek	117
Qy	126	FORYSARSLGYEFHAGHEYIYSTP-THNLJHMKLCRMKVPV-----CCASTSHSGEKPVP	180
Db	118	qflftpslsgfegrpyfviissaipdngrrscldklkvfrptnscmktigvhdr-----	173
Qy	181	LPOFTMGPNVKYNVLEDFE-----GEN-POVPKLEKISICTSPKREHLPLAVCI	228
Db	174	--vfdvndkve-nslepadtdthvsaeprsgenaatpri-----psr-----lai	217
Qy	229	AFFLMTFL 236	
Db	218	llflaml 225	
RESULT 11			
AAW00035			
ID	AAW00035 standard; Protein; 228 AA.		
XX	AAW00035;		
XX	24-OCT-1996 (first entry)		
DT	HEK4 binding protein.		
DE	HEK4		
XX	HEK4 binding protein; HEK4 receptor; EPH-like receptor;		
KW	protein tyrosine kinase; ligand; growth; differentiation; cancer;		
KW	nervous system disorder; therapy; antibody.		
XX	Homo sapiens.		
OS			
FH	Key	Location/Qualifiers	
FT	Peptide	1..119	
FT	Protein	/label= Sig_peptide	
FT		20..228	
FT		/label= Mat_protein	
PN	WO9623000-A1.		
PD	01-AUG-1996.		
XX	16-JAN-1996;	96WO-US01079.	
PF	27-JAN-1995;	95US-O379802.	
PR	(AMGE-) AMGEN INC.		
XX	Bartley TD, Fox GM;		
PI	WPI; 1996-362633/36.		
DR	N-PSDB; AAT34292.		
XX	Ligand for EPH-like receptors, partic. the HEK4 receptor - useful to		
PT	modulate growth and differentiation of, e.g. liver and kidney cells,		
PT	and to treat cancer and nervous system disorders		
XX	Claim 8; Page 39-40; 65pp; English.		
XX	Human HEK4 binding protein (HEK4 BP) (AAW00035) binds to and activates		
CC	HEK4 and ECK receptors. Its amino acid sequence was deduced from		
CC	a cDNA clone (AAT34292) isolated from a human placenta cDNA library.		
CC	Expression vectors and host cells can be used for the prodn. of		
CC	biologically active HEK BP. It is useful for modulating the growth		
CC	and/or differentiation of EPH sub-family receptor-bearing cells,		
CC	esp. in liver, kidney, lung, skin or neural tissues. It can be		
CC	used to treat neural system disorders and in the regeneration of		



	Query Match	31.0%;	Score 403;	DB 20;	Length 184;
	Best Local Similarity	50.6%;	Pred. No. 8.5e-33;		
	Matches	82;	Conservative	24; Mismatches	40; Indels
	Gaps	5;			
QY	28	ALGNRRHAVYWNSSNQHRLRE-----GYTVGVNVNDYLDIYCPHYNSGSGVGPAGPGG	81		
Dd	3	anadyavywnrnsprfqvsavdggggytvevsindyldiyphy-----gqplpaee	55		
QY	82	GAEQYVLVMSNGRYRTCNASQ-GFKRWCNRPIAPHSPIKFSKFORYSAFSLCYGFHA	140		
Dd	56	rmerillymvmngeshacdrgqrgrkfwecnrpaapggplkfsekqlftpslsgferp	115		
QY	141	GHEYIIYS-TPTNHLHWKCLRMKVFCVCCAS-TSHSGEKPVPY	180		
Dd	116	chevyvgaatccpucydrclskllkvvatctcglfccccsfaf	157		

RESULT	15
AAR23895	
ID	AAR23895 standard; Protein; 205 AA.
XX	
AC	AAR23895;
XX	
DT	06-NOV-1992 (first entry)
XX	
DE	B61 protein sequence.
XX	
KW	Early inflammatory response; marker; antibody; therapy; induction;
KW	lipopolysaccharides; cytokines; Interleukin-2; IL-2; TNF.
XX	

OS	Homo sapiens.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..18	
FT		/note= "signal peptide"	
XX			
XX	W09207094-A.		
XX			
XX	30-APR-1992.		
PD			
XX			
XX	15-OCT-1991;	9LWO-US07704.	
PF			
XX			
PR	16-OCT-1990;	9OUS-0607741.	
XX			
PA	(UNMI ) UNIV MICHIGAN.		
XX			
PI	Dixit VM;		
XX			
XX			
DR	WPI; 1992-167172/20.		
DR	N-PSDB; AAO24595.		
XX			
PT	Gene prod. used as marker to detect inflammatory response - by		
PT	detecting the B61 gene in biological fluids or by using		
PT	hybridisation probes		
XX			
PS	Claim 7; Fig 3; 39pp; English.		
XX			
CC	The B1 protein sequence was deduced from the cDNA sequence obtd. by		
CC	screening a cDNA library from human umbilical vein endothelial cells		
CC	with radiolabelled cDNA fragments derived from the 5' end of B61 DNA.		
CC	The B1 protein has 205 residues (24 kd) comprising a signal		
CC	sequence and hydrophobic N- and C- terminal regions. The B61 gene		
CC	is involved in early inflammatory response and serves as a marker.		
CC	It may be detected by probes or by antibody-based immunoassay of		
CC	biological fluids such as plasma, CSF or urine. These assays make		
CC	it possible to predict a worsening in a disease process and allow		
CC	the quantitative assessment of the magnitude of the inflammatory		
CC	response. This information will allow the earlier admin. of		
CC	appropriate therapy, thereby shortening the disease process and		
CC	limiting the patient's exposure to anti-inflammatory/immuno-		

CC limiting the patient's exposure to anti-inflammatory/immuno-  
CC suppressive therapy. B61 induction is rapid and profound, hence  
CC suppressive therapy.



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:40:20 ; Search time 1578.36 Seconds  
(without alignments)  
14904.833 Million cell updates/sec

Title: US-09-733-756-1

Perfect score: 1743

Sequence: 1 gctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hci:*
9: gb_estl:*
10: gb_est2:*
11: gb_hci:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
c 1	873.8	50.1	1008	9	AL565095	AL565095 AL565095
c 2	872.6	50.1	1030	9	AL533153	AL533153 AL533153
c 3	760.8	43.6	798	9	AL563020	AL563020 AL563020
c 4	718.8	41.2	1013	10	BF969747	BF969747 602272160
c 5	687.8	39.5	702	10	BF969747	BF969747 602272160
c 6	625.4	35.9	750	9	AL572018	AL572018 AL572018
c 7	619.6	35.5	1085	11	AK020438	AK020438 Mus muscu
c 8	615.6	35.3	684	9	AL527972	AL527972 AL527972
c 9	614.4	35.2	826	10	BF968137	BF968137 602269124
c 10	578.6	33.2	804	10	BF983120	BF983120 602305829
c 11	566	32.5	762	10	BF983120	BF983120 602305829
c 12	553.2	31.7	902	9	AL546822	AL546822 AL546822
c 13	530.2	30.4	547	10	BF99730	BF99730 602545144
c 14	439.2	25.2	496	10	BF953157	BF953157 CM3-NN118
c 15	416.8	23.9	420	9	BE138904	BE138904 xw97e11.x
c 16	413.4	23.7	435	9	AW189015	AW189015 xw98d10.x
c 17	409	23.5	535	9	AW786993	AW786993 120664 MA

c 18	401.2	23.0	406	9	AW780303	AW780303 ho10h09.x
c 19	388.8	22.3	516	10	BE663678	BE663678 147613 MA
c 20	375.4	21.5	553	10	BE663141	BE663141 126832 MA
c 21	351.4	20.2	353	9	AI272654	AI272654 q159b04.x
c 22	350.4	20.1	884	10	BF705182	BF705182 602688106
c 23	347.8	20.0	848	10	BF542148	BF542148 602068243
c 24	330.6	19.0	879	10	BF977204	BF977204 602146759
c 25	320.4	18.4	538	9	AM669878	AM669878 113619 MA
c 26	289.6	16.6	552	9	AL628407	AL628407 AL628407
c 27	279.2	16.0	654	10	BI731050	BI731050 603351667
c 28	276.2	15.8	298	9	BE048061	BE048061 tz47d12.y
c 29	274	15.7	390	9	AW123794	AW123794 UI-M-BH2.
c 30	269	15.4	523	10	BE648416	BE648416 UI-M-BH2.
c 31	257.8	14.8	460	9	AW200648	AW200648 da24c06.y
c 32	247	14.2	481	9	AA388313	AA388313 vc24h03.y
c 33	246.2	14.1	288	9	AA338749	AA338749 EST43752
c 34	245.8	14.1	788	10	BI753156	BI753156 603026017
c 35	230.8	13.2	275	10	T06173	T06173 EST04062 Fe
c 36	227.8	13.1	408	10	BF555727	BF555727 UI-R-A1-d
c 37	222.2	12.7	354	10	BG376619	BG376619 UI-R-CU0-
c 38	221.2	12.7	332	9	AI508818	AI508818 vc24h03.y
c 39	220.2	12.6	258	10	BF950361	BF950361 CM3-NN118
c 40	219.8	12.6	370	10	BE721090	BE721090 188082 MA
c 41	218.4	12.5	906	10	BF240606	BF240606 601875729
c 42	217.4	12.5	750	9	AW199923	AW199923 da24c06.x
c 43	216.8	12.4	252	9	AI863365	AI863365 tz47d12.x
c 44	215.6	12.4	364	9	AA901053	AA901053 UI-R-A1-d
c 45	206	11.8	324	10	BM386597	BM386597 UI-R-CN1-

#### ALIGNMENTS

RESULT 1  
LOCUS AL565095/c  
DEFINITION AL565095 LTI\_FL015\_Brn1 Homo sapiens cdna linear EST 16-FEB-2001  
prime, mRNA sequence.  
ACCESSION AL565095  
VERSION AL565095.1 GI:12916129  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1008)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.  
FEATURES  
Location/Qualifiers  
1..1008  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DN003YL19"  
/clone\_lib="LTI\_FL015\_Brn1"  
/sex="male"  
/tissue\_type="Adult brain"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies.  
Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"  
BASE COUNT 209 a 262 c 325 g 195 t 17 others  
ORIGIN

Query Match	50.1%	Score 873.8	DB 9	Length 1008
Best Local Similarity	94.5%	Pred. No. 1.7e-109		
Matches 943	Conservative 14	Mismatches 32	Indels 9	Gaps 5
QY 636	ccaaacgggaacacacgtgccctggccgtggcgcctcttctctatgacgtctctgg	695		
DB 991	cccaacgggaacacacgtgccctggcgcctggcgcctcttctctatgacgtctctgg	936		
QY 696	cctctagcttgcctctcctctcctctgggggggagagatgggggggcttggaaggagca	755		
DB 935	cctctctas--tctgccccctccctctgggggggagagatggggsgggc--tgaaggagca	879		
QY 756	ggagcccttggcctctccaaaggagcctagtgggcctagacccctctctccatgcta	815		
DB 878	ggga--scittggccctctccaaaggagcctagtgggcctagacccctctctccatgcta	820		
QY 816	gaagtggggcctgacatacatctgtctcggccccctcta--ccccctccccccacgtag	874		
DB 819	gaagtggggcctgacatacatctgtctcggccccctcta--ccccctccccccacgtag	760		
QY 875	ggcactgtagtggaacaaagcgggacagccatgggtcccgggcgccctgtggtctg	934		
DB 759	ggcactgtagtggaacaaagcgggacagccatgggtcccgggcgccctgtggtctg	700		
QY 935	gtaattgtgtaccaaaacttggggggccaaaaggcgagtgctcaggactccctggcccc	994		
DB 699	gtaattgtgtgtaccaaaacttggggggccaaaaggcgagtgctcaggactccctggcccc	640		
QY 995	tgttaccttccctgaactcctgttgcctctcctcttcttgcctcccccagagacatgccc	1054		
DB 639	tggtaacctttccctgaactcctgttgcctctcctcttcttgcctcccccagagacatgccc	580		
QY 1055	cccagagagaacaatcgaaagctgggagggcacccccacatgctctctccaggggcagaa	1114		
DB 579	cccagagagaacaatcgaaagctgggagggcacccccacatgctctctccaggggcagaa	520		
QY 1115	catgggagggagctagatgggcaaggggcagcactgcctgcttcttcttcccttctt	1174		
DB 519	catgggagggagctagatgggcaaggggcagcactgcctgcttcttcttcccttctt	460		
QY 1175	acagcaataagcagctcctctcccccacatccacacttccaggttgggttgattgaa	1234		
DB 459	acagcaataagcagctcctctcccccacatccacacttccaggttgggttgattgaa	400		
QY 1235	accaagttaacagtagacacccctggggggggcgagtggaacaaagga tgcacaaggggt	1294		
DB 399	accaagtgttacagtagacacacccctggggggggcgagtggaacaaagga tgcacaaggggt	340		
QY 1295	gggcattggggtgccaggcagcagtgacagactctatctctctatataatgacaga	1354		
DB 339	gggcattggggtgccaggcagcagtgacagactctatctctctatataatgacaga	280		
QY 1355	cagacagagtccttccctctttaacccctgaccttcttgacttccctctcagcttca	1414		
DB 279	cagacagagtccttccctctttaacccctgaccttcttgacttccctctcagcttca	220		
QY 1415	gaccccttccccacaggctaggccccccacacactggggagccccctgaccttcttg	1474		
DB 219	gaccccttccccacaggctaggccccccacacactggggagccccctgaccttcttg	160		
QY 1475	tcctctggaagcaggacctatgcaacgcacagacacttttggagaccgtataaacaaca	1534		
DB 159	tcctctggaagcaggacctatgcaacgcacagacacttttggagaccgtataaacaaca	100		
QY 1535	gcgccccctcccttccagcctgagccggaaacacatccacaggaccttgcctgctcac	1594		
DB 99	gcgcccccttsswtctyagccttgaggccgggaagsgttctcccaggaccttgcctgctcac	40		
QY 1595	cctatgtgtcccaactatctctctgggctttttca 1632			
DB 39	ccatgtgtgtgtcccaactatctctctgggctttttcna 2			

RESULT 2  
AL533153

LOCUS	AL533153	1030 bp	mRNA	linear	EST 13-FEB-2001
DEFINITION	AL533153	LTI_FL015_Brr1	Homo sapiens	cDNA clone CS0DN003YL19 5	
		prime, mRNA sequence.			
ACCESSION	AL533153				
VERSION	AL533153.1	GI:12796646			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1030)				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.				

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DN003YL19"  
 /clone\_lib="LTI\_FL015\_Brl"  
 /sex="male"  
 /tissue\_type="Adult brain"  
 /note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"  
 183 a 333 c 338 g 173 t 3 others  
 BASE COUNT  
 ORIGIN

Query Match	50.1%	Score 872.6;	DB 9;
Best Local Similarity	97.7%	Pred. No. 2.4e-109;	
Matches 945;	Conservative	3;	Mismatches 12;
			Indels 7;
			Gaps 6;

1 gctgctgctgctgctgctgcgtgccgtgcgcgtgctgccgtgctgcccgaaggcc 60  
Db |  
66 GCTGCTGCTGCTGCTGCTGCTGCCGTGCCGTGCTGCCGTGCTGCCCAAGGGCC 125  
QY |  
61 cggagggcgctggaaaaccgcacatgcggtgactggaaacagctccaaccagcactgcg 120  
Db |  
126 CGGA-GGGCGCTGGAAACCGGCATGCGGTGACTTGAACAGCTCCAACACGACCTGG 184  
QY |  
121 gcgagagggtcacccgtgcagtgcaacgtgaacqactatctggattattactccccga 180  
Db |  
185 GCGAGGGGCTACACCGTCAGTGTAACGTGAACGACTATCTGGATTATTTACTGCCCGCA 244  
QY |  
181 ctacaacagctcgggggtggccccggggcgggagccggccgcgagggcgggcgagagca 240  
Db |  
245 CTACAACAGCTCGGGGTGGGGCCCCGGGGCGGACC GGCGGCCGGAGCGGGCAGAGCA 304  
QY |  
241 gtacgtgctgtacatgggtgagccgcaacgcgttacgcgacctgcacgcgcagccaggctt 300  
Ddb |||  
305 GTACGTGCTGTACATGGTGTAGCGCGAAGCGGTACCGGCACCTGTCAACGCGCAGCAGGGGTT 364  
QY |  
301 caagcgctgggagtgcaaccggcgccagcccccgcacagccccatacagtctcggagaa 360  
Ddb |||  
365 CAAGCGCTGGGANGTC AACCGGCCGACGCCCGGCACAGCCCCTCAAGTTCTCGGAAA 424  
QY |  
361 gtccacgcgtacagcgcccttcctct-gggctacgagttccacgcggcgccacgagtact 419

```
Db 425 GTTCCAGCGCTACAGSCCTTCTCTGGGCTACGAGTCCACGCCGCGCCAGGACTACT 484
QY 420 actacatctccagcccaactcaactcactcactggaagtgtctgagatgaagggttctcg 479
Db 485 ACTACATCTCCAGCCCACTCAACACTGCACCTGGAAGTCTCTGAGGATGAAGGTGTTCG 544
QY 480 tetgtgcgcctccacatgcactcc- ggggagaagccggtcccaactctccccagttc 538
Db 545 TCTGTCTGCCCTCCACATCGCACTCGGGGAGAGCCGGTCCCACTCTCCCCAGTTC 604
QY 539 accatgggcccacatgtgaagatcaactgtctggaagactttgagggagagaaacctcag 598
Db 605 ACATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGAGAGAACCCTCAG 664
QY 599 gtgcccagaacttgagaagatcagcgagaccagcccaaacggaagacacactgcccctg 658
Db 665 GTGCCCAAGCTTGAGAGAGCATCAGCGGACCGCCCAACGGGAACACCTGGCCCTG 724
QY 659 gcctgtggcactgccttctctcatgacgtttcttgccctcctagctctgccccctccc 718
Db 725 GCCGTGGGATCGCCTCTCTCTCATGACGTTCTTGGCCTCTAGCTCTGCCCTCCCTCCC 784
QY 719 tgggggggagagatggggcggggttggaaagacaggagagccttggcctctccaagg 778
Db 785 TGGGGGGGAGAGATGGGGCGGGGCTTGAAGGAGCAGGGAGCCCTTGGCCCTCTCCAAG 844
QY 779 gaagcctagtgggcctagaccctcctccatggctagaagtggggcctgaaccatacat 838
Db 845 GAAGCCTAGTGGGCTAGACCCCTCTCCCATGGCTAGAAAGTGGGCTGCACCATACAT 904
QY 839 ctgtatcgccctctacccttccccacgttagggcactgtagtgaccacacagg 898
Db 905 CTGTGTCCGCCCTCTTACCCCTTCCGCCCTAGTGGGCTCTGTTGTTTACAAATMTGGG 961
QY 899 ggacagcctagtggtcccgggcggtctgtgctctgtgtaatttggtaaccaacttggg 958
Db 962 GGACAG-CATGGTCTCCCGGGGCGCTTGTGGCTCTGGTAAATGTTGTTTACAAATMTGGG 1020
QY 959 ggcceaaa 965
Db 1021 GCCAAA 1027

RESULT 3
AL563020/c
LOCUS AL563020 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YP19 3
DEFINITION prime, mRNA sequence.
ACCESSION AL563020
VERSION AL563020.1 GI:12912020
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 798)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YP19"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

FEATURES
source
```

/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 164 a 203 c 257 g 165 t 9 others

Query Match 43.6%; Score 760.8; DB 9; Length 798;

Best Local Similarity 97.9%; Pred. No. 3.7e-94;

Matches 782; Conservative 8; Mismatches 6; Indels 3; Gaps 2;

QY 745 tgaaggagcagggagcctttggcctctccaagggaagcctagtggcctagacccctcc 804

Db 796 TKGAGGAGCAGGGAGC--TTTGCTCTCCAAGGAAGCCTAGTGGCCCTAGACCCCTCC 739

QY 805 tcccatgctagaaagcgggagcctgcaccatacatgtgtccgccccctctacccttcc 864

Db 738 TCCCATGGCTAGAGT-GGGCCTGCACCATACATCTGTCTCGGCCCTTATACCCCTTCC 680

QY 865 ccccaactgagggcactgtaggaccagcagggagacatggtgccggcgccct 924

Db 679 CCCACAGTAGGGCAGCTAGTGGACCAAGCAGCGGACAGCATGGGTCCCGGGCGGCT 620

QY 925 tctgctctgtaattgtttgtglaacaaacttgggggcccacaaagggcagtgctcaggact 984

Db 619 TGTGGCTCTGTAATGTTTGTGTACAACTTTGGGGGCCCAAAAGGCGAGCTCAGGACT 560

QY 985 ccttggcccccttggactcttccctgactcctgtggcctctccctttatcccccaaga 1044

Db 559 CCCTGGCCCTGGTACCTTCCCTKACTCTCTGTGGTCTCTCCCTTTGTCCCCGAGAGA 500

QY 1045 gacatatgccccagagagaaatcgaagcgtgggagggcaccaccaatgtctctctcc 1104

Db 499 GACATATGCCCCAGAGAGAGCAAAATCGAAGCGTGGGAGGACCCCACTTCTCTCTCC 440

QY 1105 agggcagaaactggggagggagtagatgggcaaggggcagcactgctctcttct 1164

Db 439 AGGGCAGAAATGGGAGGGGACTAGATGGCAAGGGGAGCAGCTGCTCTGCTCTCT 380

QY 1165 tccctgtttacagaataagcagctctctctccccactccccacttccagattgtgt 1224

Db 379 TCCCTCTGTTTACAGCAANAAGCAGCTCTCTCTCCCAACTCCCACTTCCAGGATTGTGT 320

QY 1225 ttgattgaacaaactttcaagtagacaccccttggggggcgggcagtgagacaagat 1284

Db 319 TTGATTGAAACCAAGTTTACAGTACACACCCCTGGGGGGGGGGGAGTGTGACAGAT 260

QY 1285 gccaaagggtggcattggggtgccaggcagcattgtacagactctatatctatatat 1344

Db 259 GGCAAGGGGTGKGGATTGGGGTGCCAGCAGGAGCATGTACAGACTCTATATCTCTATAT 200

QY 1345 aatgtacagacagacagagtccttccctcttttaacccccctgaaccttcttgactcccc 1404

Db 199 AATGAACAGACAGACAGAGTCCCTTCCMTCTTTAAACCCCTTGACCTTCTTGTGACTTCCC 140

QY 1405 ttcagcttcagacccttccccaccagctaggccccccacacctggggaccccttggc 1464

Db 139 TTCAGCTTCAGACCCCTTCCCCAGGCTAGGGCCCCCAGACCTGGGGGAGCCCTTGGC 80

QY 1465 cctcttttctctctgtgaagacagacactatgcaacgcacacacacttttggagacc 1524

Db 79 CCCTCTTTTCTCTGTGAAGACAGGACCTATGCAACGACGACAGACACTTTTGGAGACCG 20

QY 1525 taaaacaacagcccccct 1543

Db 19 TAAWACAACAGCGCCCT 1



## ORIGIN

Query Match 39.5%; Score 687.8; DB 10; Length 702;  
Best Local Similarity 99.6%; Pred. No. 3e-84;  
Matches 700; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 555 tgaagatcaacgtgctggaagactttgagggagagaaacacctcaggtgcccagcttgaga 614  
DB 1 TGAAGATCAACGTGCTGGAAGACTTTGAGGGAGAGAACCTCAGGTGCCACAGCTTGAGA 60  
QY 615 agagcatcagcgggacaccccccaacgggaacacctgccccctggcgtgggcatcgct 674  
DB 61 AGAGCATCAAGCGGACACCCCAACAGGGAACACCTGCCCTGGCGGTGGCATCGCT 120  
QY 675 tctctcatcagcttcttgctctctagctctgccccctccccctggggggagagatg 734  
DB 121 TCTTCTCATGAGTCTTGGCTCTCTAGCTCTGCCCTCCCTGGGGGGGAGAGATG 180  
QY 735 gggcggggttggaagagcaggagacctttggctctcacaagggaagcctagtggcct 794  
DB 181 GGGCGGGGTGGAAGAGCAGGGAGCCCTTTGGCTCTCCAAGGAAGCCTAGTGGCCT 240  
QY 795 agacctctctccatcgctagagtgaggcctgacacatcatctgtccgccccctc 854  
DB 241 AGACCTCTCTCCATGGCTAGAGTGGGCTGTGACCATATCTGTGTCGCCCTCTC 300  
QY 855 taacctctcccccaagtagggcactgtagtggacaagcagcgggacagcctatgggtcc 914  
DB 301 TAGCCCTTCCCTCCACGTAGGCACTGTAGTGACCAAGCAGCGGGACAGCCATGGTCC 360  
QY 915 cggcgcccttgactctggttaattgtgtacaaacttgggggcaaaaaggagcagt 974  
DB 361 CGGCGGCTTGTGCTGCTGTAATGTTGGTACCAACTTGGGGGCCAAAAGGCGCAGT 420  
QY 975 gctcaggactcctggccctgtgacctttccctgactcctggtgctcctctctcttgc 1034  
DB 421 GCTCAGGACTCCTGGCCCTGGTACCTTTCCCTGACTCTCTGGTCCCTCTCCCTTTGTC 480  
QY 1035 ccccaagagacatagctccccagagagagcaaatcgaagcgtggggagcaccctatt 1094  
DB 481 CCCCCAGAGACATATGCCCCAGAGAGAGCAATCGAAGCGTGGGAGGACCCCTCAT 540  
QY 1095 gctcctccagggcgaacatggggagggactagatggcgaaggcagcactgct 1154  
DB 541 GCTCTCTCCAGGCGCAACATGGGAGGGGACTAGATGGCAAGGGCGACACTGCT 600  
QY 1155 gctgctctctccctgtttacagcaataagcagctcctctcccccaactccacttcca 1214  
DB 601 GCTGTCTCTCCCTGTTTACAGCAATAAGCAGCTCTCTCTCCCTCCCA-TCCCCACTTCA 659  
QY 1215 ggtatggtttggttgaacaaagtttaagtagacacc 1257  
DB 660 GGATCGTGTTCGGATTGAACCAAGTTTACAGTAGACACC 702

RESULT 6  
AL572018/c  
LOCUS  
DEFINITION AL572018 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1026Y124 3  
prime, mRNA sequence.

ACCESSION AL572018

VERSION AL572018.1 GI:12929883

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 750)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 750

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0D1026Y124"

/clone\_lib="LTI\_NFL006\_PL2"

/tissue\_type="placenta"

/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies. Contact : Feng Liang Life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 137 a 199 c 241 g 153 t 20 others

ORIGIN

Query Match 35.9%; Score 625.4; DB 9; Length 750;

Best Local Similarity 94.3%; Pred. No. 8.4e-76;

Matches 706; Conservative 16; Mismatches 18; Indels 9; Gaps 7;

QY 802 tctctccatgctagaagtgggctgcacacatcatctgtgtccgccccctctaccctc 861

DB 750 TCTTCCCATAGCTAGAGTGGGCTGCACCATACATCTGTGTCCGCCCTCTACCCCT 691

QY 862 tcccccacagtagggcactgtagtgaccacagcgggagcagcctatgggtcccgggcg 921

DB 690 YCCCCCACSTAGGCACTGTAGTGGACCAAGCAGCGGDACAGCCATGGGTCCCGGGCG 631

QY 922 ccttgaggctgttaattgttgtaacaaacttgggggcaaaa-aaggcagtgctcag 980

DB 630 CTTTGTGGCTCTGTAATGTTGTGTACCAAACTTGGGGCCAAAAGGCGAGTGTCTAG 571

QY 981 gactccttgccctggtgacctttccctgactcctggtgacctcctcttctgccccca 1040

DB 570 GACTCCCTGGCCCTGGTACTCTTCCCTRACTCT-GTGCTCTCTCCCTTTGTCCCTCCA 512

QY 1041 gagagacatatgccccagagagagcaaatcgaagcgtggggagcaccctattgctctc 1100

DB 511 GAGAGACATATGCCCCAGAGAGAGCAATCGAAGCGTGGGAGGACCCCTCATTTGCTCTC 452

QY 1101 ctccagggcagaaacatggggaggggactagatgggcaaggcagcactgctcgtcgt 1160

DB 451 CTCAGGGGCGAGACATGGGAGGGGACTAGATGGGCAAGGGGAGGACTGCCCKGTGCT 392

QY 1161 tcttccccctgtttacagcaataagcagctcctcctcccccaactccccctcaggattg 1220

DB 391 TCCTTCCCTGTTTACAGCAATAAGCAGCTCCCTCCCTCCCMAGCCCCCTTACCAGGATGG 332

QY 1221 tggttggattgaacacaaagtttacagtagacacccctggggggggcagcagtgagacaa 1280

DB 331 TGGTTTGGATGGAACCAAGTTACAAAGTAGACACCCCTGGGGGGGGGCGAGTGAGCAA 272

QY 1281 ggatgccaagggtggcgattggggtgccaggcagcagcatgtacagactctatatct-cta 1339

DB 271 GGATGGCAAGGGTGGGCATTGGGGTGCCAGGCGAGCATGTACAGACTCTATAGTCTTA 212

QY 1340 tatataatgtacagacagacagagctcctctctctttaaacccttgactcttcttgaact 1399

DB 211 TATATAATGTACAGACAGACAGAGGCGCTTCCCTCTTTTAACTCCCTGAGCTKWCYTGACT 152

QY 1400 tc-cccttcagcttcagaccccttccccaccagctagcccccca-cacctggggagacc 1457

DB 151 TCGCCCTYAGGCCCGCAGACCCCTTCCCCACAGGCTTAGGCCCCCCCCCCAGACCTTGGGGGACC 92

QY 1458 ccttgggccctcttttctct-ctgtggaagacagagccta-tgcaacgcacagacact 1513





```

Db 541 GAAGATMACGCTGCTGGANAGACTTTGAGGAGAGAGAACCTCAGGTGCTCCCAAGCTTGAGAA 600
Qy 616 gagcatcagggaccagcccaacgggaacacacgtgcccctggccgtggcgtgagcatgcctt 675
Db 601 GAGCATCAGCGGACAGCCCAACAGGGAACACCTGCCCCCTGGCCGTGGCATGCCTT 560
Qy 676 ctccctcatgagcttctggcc 697
Db 661 CTTCCCTCMCGACGTATCTTGGC 682

RESULT 9
BF968137 826 bp mRNA linear EST 22-JAN-2001
LOCUS 602269124F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4357493 5',
DEFINITION mRNA sequence.
ACCESSION BF968137
VERSION BF968137.1 GI:12335352
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14W993 row: p column: 06
High quality sequence stop: 769.
Location/Qualifiers
1..826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4357493"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
Not1; Site_2: SalI; Cloned unidirectionally; oligo-dr
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 180 a 260 c 208 g 178 t
ORIGIN

Query Match 35.2%; Score 614.4; DB 10; Length 826;
Best Local Similarity 89.4%; Pred. No. 2.5e-74;
Matches 735; Conservative 0; Mismatches 66; Indels 21; Gaps 6;

Qy 864 ccccccacgtagggcactgtagtgcacaaagcacggggacagccatgggtccggcgcc 923
Db 1 CCCCCACGTAGGGCACTGTAGTGGACCAAGCACGGGACAGCCATGGTCCCGGGCGGCC 60

Qy 924 ttgtggtctgttaattttgtttaccacaaacttggggcccaaaagggcagtgctcaggac 983
Db 61 TTGTGGCTCTGGTAATGTTTGGTACCAAACTTGGGGGCCAAAAGAGGCGAGTGTCTCAGGAC 120

Qy 984 tccctggccctggctaccttccctgactcctgtgctcctctccttcttgcctcccccagag 1043
Db 121 TCCCTGGCCCTGGTACTTCCCTGACTGCTGCTGGTGGCCCTCTCCCTTTTGCCCCCAGAG 180

Qy 1044 agacatatgccccagagagcaaatcgaagcgtggagggaccccccttgcctctc 1103

```

```

Db 181 AGACATATGCCCCCAGAGAGAGAAATCGAAGCGTGGGAGGCACCCCATTCCTCCTC 240
Qy 1104 cagggcagaacatggggaggggactagatgggcaaggggcagcactgcctgctgtcc 1163
Db 241 CAGGGCGAAGACATGGGGAGGGGACTAGATGGCAAGGGGAGCAGCTGCCTGCTGCTTCC 300
Qy 1164 ttccctgtttacagcaataagcagctcctcctcccccactcccacttcccaggatgtg 1223
Db 301 TTCCCTGTTTACAGCAATAAGCAGCTCCTCCTCCCCACTCCCACTCCAGGATTGTGG 360
Qy 1224 ttgagattgaacaaagtttacaagtagacacccctggggggcggcagtggaagga 1283
Db 361 TTTGGATTGAACCAAGTTTAAAGTAGACACCCCTGGGGGGGGGCGGAGTGGACAAGGA 420
Qy 1284 tgcaaggggtggcattgggtgcccagcagcatgtacagactctatctctatata 1343
Db 421 TGGCAAGGGGTGGGCATTTGGGGTGGCAGGAGCATGTACAGACTCTATCTCTATATA 480
Qy 1344 taatgtacagacagagagtccttccctctttaaacccttgacctttcttgcctccc 1403
Db 481 TAATGTACAGACAGACAGAGTCCCTTCCCTTTAAACCCCTGACCTTCTTGTGACTTCCC 540
Qy 1404 cttcagcttcaga--cccttcccccacagcagctagcccccacacacacacacacccct 1461
Db 541 CTTGAGCTTACAGACCCCTTCCCTCCACAGGCTAGG--CCCCACAACTGGGGGACCCCT 599
Qy 1462 gggccctcttctgtctgtggaagacagcagctatgcaacgcacaga-----cacttt 1516
Db 600 GGCCCATCTTTGCTTCTGTGAAGAAGCAGGACCTATGCTGAGGCGCACAGACCACTTTG 659
Qy 1517 ggaacacgtaaaaaacacagcgcctccctccctcagcctgagccggggaacacacac 1576
Db 660 GAACACCGTAAACAAACACAGCCCTCCCTTTTCAGCCCTGAGCCGGGAA---CATCTC 715
Qy 1577 ggaccttgcctgcacacctatgtgtccacacatactcctcctgggaccttttcaagtgc 1636
Db 716 CAGGACTTGCTGCTCACCCTATGTGTGTCACATATCTCC-----GGCTTTTCAAGTGC 769
Qy 1637 ttggctgtgacttctatactgctcttctttagtcttaaaaaa 1678
Db 770 TTGCTGTG---AATTCATACCTGCTCTCTTGTAAACAAAAA 808

RESULT 10
BF983120 804 bp mRNA linear EST 23-JAN-2001
LOCUS 602305829F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4397263 5',
DEFINITION mRNA sequence.
ACCESSION BF983120
VERSION BF983120.1 GI:12386017
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14W10097 row: i column: 08
High quality sequence stop: 766.
Location/Qualifiers
1..804
/organism="Homo sapiens"
source

```

LOCUS	CG698896	762 bp	mRNA	linear	EST 07-MAY-2001
DEFINITION	602703313f1 NCI_CGAP_skn3 Homo sapiens cDNA clone IMAGE:4800482 5', mRNA sequence.				
ACCESSION	CG698896				
VERSION	CG698896.1	GI:13966642			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 762)				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 762)	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a>	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact:	Robert Strausberg, Ph.D.			
Email:	cgapbs@mail.nih.gov			
Tissue	cgapbs-r@mail.nih.gov			
cDNA Library	Preparation: James Cleaver, M.D.			
cDNA Library	Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA			
Sequencing	by: Incyte Genomics, Inc.			
Clone distribution:	MGC clone distribution information can be			
found through	the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov				
Plate:	LLAM10691 row: j column: 03			
High quality	sequence stop: 762.			
Location/Qualifiers				
1..762				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:4800482"				
/clone_lib="NCI CGAP_Skn3"				
/lab_host="DH10B (T1 phage-resistant)"				
/note="organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;				
Site_2: SalI; Cloned unidirectionally. Primer: Oligo.dT.				
Average insert size 1.5Kb. Library constructed by Life				
Technologies. Note: this is a NCI CGAP Library."				
BASE COUNT	149 a	235 c	216 g	161 t
ORIGIN				
Query Match	32.5%	Score 566;	DB 10;	Length 762;
Best Local Similarity	94.8%;	Pred. No. 9.4e-66;		
Matches 704;	Conservative 0;	Mismatches 26;	Indels 13;	Gaps 11;
659	gcgcgtggcgatcgcttctctcatgacgtttcttgccctctagctctgccccctccc	718		
3	GCCTGTCTCATGAGTCTTTGG-CTCCTAGCTCTGCCCTCCCC	60		
719	tggggggagagatggggcggttggaaggagcaggagcctttggcctctccaaag	778		
61	TGGGGGGGAGAGATGGGGGGGGCTTGGAGGAGCAGG--AGGCTTGGCTCTCCAAAG	118		
779	gaagcctagtggcctagacccctctccatggctagaaagtggggcctgaccatacat	838		
119	GAAGCCTAGTGGGCTAGACCCCTCTCCATGGCTAGAGTGGGGCTGCACCATACAT	178		
839	ctgtgtccgcgccctctacccttccccccacgtagggcactgtagtggaccagacgg	898		
179	CTGTGTCCG-CCCCCTTACCCTTTCCCCCAGCTAGGGCAGCTAGTGGACCAAGCAGG	237		
899	ggacagccatgggtccgcggcgccctgtgctctggttaattggtaccacaaactggg	958		
238	GCACAGCATGGTCCCGGGCGG-CTTGTGGCTCTGGTAAATGTTGGTACCAAACTGGG	296		
959	ggccaaaaagggcagtgctcaggactccctggccccctgggtaccttccctgactccgtt	1018		
297	GGCCAAAAGGGCAGTGCTCAGGACTCCCTGG-CCCTGGTACCTTTCCCTGACTCCTGGT	355		
1019	gccctctcccttggccccccagagacacatgtccccccagagagacaaatcgagcgt	1078		
356	GCCCTCTCCCTTTGTCCTCCCGCAGAGACATATGCCCCCGCAGAGACAAATCGAACGCT	415		

Db 416 -GGAGGCAACCCCAATGCTCTCCCTCCA -GGGCAACAATGGGAGGGAGCTAGATGGGCA 473  
 Qy 1139 agggcagcactgcctgctctcccttccctgtttacagcaataagcagctcctctcc 1158  
 Db 474 AGGGGCGAGCACTGCTGCTCTCCCTCTGTTTACAGCAATAGCACGCTCTCTCTCC 533  
 Qy 1199 cccactccactccagcattggtgttgattgaaacaaagtttacaagtacaccccc 1258  
 Db 534 CCCA-TCGCCACTTCAGAGT--GTGGTTGGATTGAACCAAGTTTACAAGTAGACACCCC 590  
 Qy 1259 tggggggcgagcagtgagcaagatcccaaggggtggcattgggtgcccaggcagga 1318  
 Db 591 TGGGGGGCGGCGAGTGGACAAGGATGCAAGGGGTGGGCANT-GGGTGGCAGGCA 649  
 Qy 1319 tgtacagactctatctctctatatataatgtacagacagagctcccttccctcttta 1378  
 Db 650 TGTACAGACTCTATATCTATATATAATGTACAGACAGACAGAGTCTTCCCTCTTTAA 709  
 Qy 1379 acccctgaccccttcttgaactt 1401  
 Db 710 CCCGGGACTCTTGATTCCTTC 732

## RESULT 12

AL546822 902 bp mRNA linear EST 16-FEB-2001  
 LOCUS AL546822 LTI\_NFL006\_PL2 Homo sapiens CDNA clone CS0D1026Y124 5  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL546822  
 VERSION AL546822.1 GI:12880311  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1. (bases 1 to 902)  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

## FEATURES

source  
 1. 902  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="LTI\_NFL006\_PL2"  
 /tissue\_type="placenta"  
 /note="Vector: pCMVSPORT 6; site\_1: NotI; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed by  
 Life Technologies. Contact: Feng Liang Life Technologies,  
 a division of Invitrogen 9800 Medical Center Drive  
 Rockville, Maryland 20850, USA Fax: (1) 301 610 8371  
 Email: fliang@lifetech.com URL:  
 http://fulllength.invitrogen.com"  
 BASE COUNT 161 a 297 c 278 g 156 t 10 others  
 ORIGIN

Query Match 31.7% Score 553.2; DB 9; Length 902;  
 Best Local Similarity 91.1%; Pred. No. 4.9e-66;  
 Matches 749; Conservative 8; Mismatches 37; Indels 28; Gaps 15;  
 Qy 105 ccaacacagcactgcggcgaggaggtacacccgtgacgtgaa-----cga 156  
 Db 67 CCCACAGCCTGGGGCGGAGCGGCTACACCGTGCAGTGAACGTGCAGAGAGCTCAC 126

Qy 157 ctatctgatatattactgcccgcactacaacagctcgggggtggccccggggcg--g 213  
 Db 127 TCGTGGATATGTTAGCTTGCCTCGCACTACACAGCTCGGGGTGGCCCCGGGGCGAC 186  
 Qy 214 accggccggggagggcgggcaga-gcagttacg-tgctgtac-atggtgagcgcaacgg 270  
 Db 187 CGCGCGCGCGGAGGCGGGCAGACGCGAGTAGCGTGTACGATGCTGAGCGCGCAACGG 246  
 Qy 271 ctaccgcactgcgaacgcccagcagcgcttcaagcgtggtgagtgcaacggcgcaacgc 330  
 Db 247 CTACCGCACCTGCAACGCCAGCCAGGGCTTCAAGCGCTGGGAGTGC AACCGCGCGCACGC 306  
 Qy 331 ccgcacagccccat-caagtctcggagagttccagcgctaca--gcgcctctctct 387  
 Db 307 CCCGCACAGCCCCATGCAAGTTCTCGGAGAAGTTCCAGCGCTACAGCGCGCTTCTCTCT 366  
 Qy 388 gggctcagagttccacgcgcggccacagtgactactactacatccacgcccactcacaacct 447  
 Db 367 GGGCTACGAGTTCCACGCGCGCCACGAGTACTACTACATCTCCACGCCCTACTCACAACCT 426  
 Qy 448 gcactggaagtgtctgagatgaagtggtctgctgctgcgctccacatcgcaactcgg 507  
 Db 427 GCACCTGGAAGTGTCTCAGGATGAAGGTGTCTGCTGCTGCGCTCCACATCGCCTCCGG 486  
 Qy 508 g-gagaagcggtcccccactctccccagttccaccatggggccccaatgtgaagatcaacg 566  
 Db 487 GCGAGAAGCGGCTCCCTCACTCTCCCGAGTTCACCATGGGCCCCCAATGTGAAGATCAACG 546  
 Qy 567 tgcggaagacttggggagagaccctcaggtgcccaagcttgagaagacat--cag 624  
 Db 547 TGTGGAAGACTTTGAGGAGAGAACCTTCAGTGCCCAAGCTTGAGAGAGCATCGACGG 606  
 Qy 625 cgggaccagccccaaacgggaacacctgcccctggcgtggcgatcgcccttctctctcat 684  
 Db 607 CGGGACAGCCCAACAGCGGAACACCTGCCTCGCGCTGGCGTGGCATCGCCTTCTTCTCAT 666  
 Qy 685 gaagcttggcctcactagctctgcccctccctc--ggggggggagagatggggcgggc 743  
 Db 667 GAGCTTCTTGGCTCTCTAGCTCTGCCCTCCCGGGGGGGGAGAGATGGGGGGGGG 726  
 Qy 744 ttggaagagcagggagcctttggcctctccaaagggaagcctagtgggcctagacccctc 803  
 Db 727 TTGGAAGGAGCAGGAGGCGCTTTGGCTCTCCAA-GGAACCTAGTGGGCTAGA-CCCTC 784  
 Qy 804 ctcccatggtagaagtgggcgctgcacacatcatctgtgtccgcccctctaccccttc 863  
 Db 785 CTCCTCATGGTAGAAGTGGGCTGCACCATACATCTGTGTCCGCC--CTTACCCCT 841  
 Qy 864 ccccccacgtaggcactgtagtggaacacgacgagggacgc 905  
 Db 842 CCCCCACGTA-GGCACCTGTAGTGA-CAAGCACGGGGACABC 881

## RESULT 13

BG499730 547 bp mRNA linear EST 27-MAR-2001  
 LOCUS BG499730 NIH\_MGC\_60 Homo sapiens CDNA clone IMAGE:4667555 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG499730  
 VERSION BG499730.1 GI:13461247  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1. (bases 1 to 547)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cyapbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CMI476 row: 9 column: 12  
 High quality sequence stop: 537.

FEATURES  
 source  
 Location/Qualifiers  
 1..547  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4667555"  
 /clone\_lib="NIH\_MGC\_60"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="pH10B (T1 phage-resistant)"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
 Site: 1: Sfll (ggccgctcgccc); Site: 2: Sfll (ggccatgatggcc  
 ); Double-stranded cDNA was prepared from cell line RNA.  
 5' and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CAGGCCATATGGCC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCCGAGCGGCCGACATG-dt(30)BN-3'  
 (where B = A, C, G, or T). Average  
 insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 112 a 171 c 146 g 118 t  
 ORIGIN

Query Match 30.4%; Score 530.2; DB 10; Length 547;  
 Best Local Similarity 99.3%; Pred. No. 7.3e-63;  
 Matches 543; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 898 gggacagcatgggtcccgggcgcttggtgctctggaatgttgatcaaaacttgg 957  
 Db 1 GGGACAGCATGGGTCCCGGGCGCTTGTGGCTCTGGTAATGTTGGTACCAACTTGG 60  
 QY 958 gggcaaaaggagtgctcaggactccctggcccttggtaccttccctgactcctgg 1017  
 Db 61 GGGCAAAAGGAGTGTCTAGGACTCCCTGGCCCTGGTACCTTTCCTGACTCCTGG 120  
 QY 1018 tgcctctctcttggccccccagagacatatccccccagagagagacaaatgaagcg 1077  
 Db 121 TGCCCTCTCCCTTTGTCCCCCAGAGACATATGCCCCAGAGAGAGCAATCGAAGCG 180  
 QY 1078 tgggggcaacccccattgtctctccaggggagacacatgggggggactagatggcg 1137  
 Db 181 TGGGAGGACACCCCATTTGCTCTCCAGGAGCAATCGGAGGAGGACTAGATGGCG 240  
 QY 1138 aaggggagcagctgctgctctctccctgtttacagaataaagacagtcctctc 1197  
 Db 241 AAGGGGAGCACTGCTGCTCTCTCCCTGTGTACGAATAAAGCAGTCTCCTC 300  
 QY 1198 cccactccactccaggattggttgattgaacacaaagtttacaagtagacaccc 1257  
 Db 301 CCCCACCTCCCACTTCCAGGATTGGTTGGATTGAACACCAAGTTTACAGTAGACACCC 360  
 QY 1258 ctggggggcgagctgacagagatgccaaaggggtgggcatgggggtgccagcgagcg 1317  
 Db 361 CTGGGGGGGGGGAGTGGAGCAAGGATGGCAAGGGGTGGGCATTGGGGTGGCAGGCGAGC 420  
 QY 1318 atgtacagactctatatctctatatataatgtacagacagacagagtccttccctcttt 1377  
 Db 421 ATGTACAGACTTATATCTCTATATAATGTACAGACAGACAGATCCCTTCTCCTTTT 480  
 QY 1378 aacccctgaactttcttgacttcccttgcagtcagaa-cccccttccccaccagggtag 1436  
 Db 481 AACCCCTGACCTTCTTGTGACTTCCCTTTCAGCTTCAGACCCCTTCCCCACCAAGGTAA 540  
 QY 1437 gcccccc 1443  
 | | | | |

Db 541 GGCCCCC 547

RESULT 14  
 LOCUS BF953157/c  
 DEFINITION CM3-NN1187-111100-478-a06 NN1187 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF953157  
 VERSION BF953157.1 GI:12370432  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 496)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-NN1187-  
 111100-478-a06&t3=2000-11-11&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 19  
 High quality sequence stop: 495.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..496  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NN1187"  
 /dev\_stage="Adult"  
 /note="Organ: nervous normal; Vector: puc18; Site: 1: SmaI;  
 Site: 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 94 a 144 c 162 g 96 t  
 ORIGIN

Query Match 25.2%; Score 439.2; DB 10; Length 496;  
 Best Local Similarity 97.9%; Pred. No. 1.6e-50;  
 Matches 466; Conservative 0; Mismatches 8; Indels 2; Gaps 2;  
 QY 443 aacctgcactggaagtgtctgagatgaaggtgtctgtctgctgcctcc-acatcgca 501  
 Db 496 AACCTGCACGTGGAAGTCTCTGAGGATGAAGGTGTCTTCTGCTGCGCTTCCAAATCGCA 437  
 QY 502 ctccggggagaagccggtccccactctccccaggttccaccatggggcccaatgtgaagat 561  
 Db 436 CTCGGGGAGAAGCGGTCCGCCACTCTCCGCCAGTTCCACCATGGGCCCAATATGAAGAT 377  
 QY 562 caactgtctggaagactttgagagagaacccctcaagtcagtcagcttgagaagacat 621  
 Db 376 CAACGTCTGGAAGACTTTGAGGAGAGAACCCCTCAGGTGCCCAAGCTTGAGAGAGCAT 317



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:43:30 ; Search time 216.15 Seconds  
(without alignments)  
13844.931 Million cell updates/sec

Title: US-09-733-756-1  
Perfect score: 1743  
Sequence: 1 gctgctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	949.2	54.5	1070	16	AAT03883
2	917.8	52.7	1037	16	AAT05887
3	447.8	25.7	715	21	AAAO1637
4	252.2	14.5	257	21	AAAO0162
5	205.8	11.8	358	21	AAC03424
6	149.2	8.6	642	20	AAK32767
7	149.2	8.6	783	17	AAT15009
8	149.2	8.6	783	19	AAV42927
9	139.6	8.0	1809	17	AAT15008
					Eph transmembrane
					Human hek-L protei
					Human colon cancer
					Human secreted pro
					Human LERK-6 poly
					Chicken elf-1 cDNA
					cDNA encoding an a
					Mouse Elf-1 cDNA.

10	139.6	8.0	1809	19	AAV42926	cDNA encoding a ma
11	135.4	7.8	314	17	AAT14010	LERK-6 exon. Homo
12	135.4	7.8	314	20	AAK32766	Exon sequence of h
13	128.8	7.4	555	17	AAT32700	Lerk-6 coding sequ
14	128.8	7.4	555	17	AAT14009	LERK-6 coding sequ
15	128.8	7.4	555	20	AAK32761	Murine LERK-6 poly
16	116.4	6.7	1438	13	AAQ24595	B61 cDNA sequence.
17	116.4	6.7	1480	15	AAQ63770	B61 gene. Homo sa
18	116.4	6.7	1496	22	AAC21387	Human cDNA sequenc
19	116.4	6.7	1496	22	AAC90574	Human PRO202 cDNA.
20	116.4	6.7	1496	22	AAC91576	Human PRO202 cDNA.
21	116.4	6.7	1527	21	AAC98823	Human pancreatic c
22	114.8	6.6	439	22	AAH81578	Human differential
23	102	5.9	687	17	AAT32699	Lerk-7 coding sequ
24	102	5.9	1176	23	AAS92024	DNA encoding novel
25	102	5.9	1574	23	AAS92779	DNA encoding novel
26	102	5.9	1728	17	AAT34292	HEK4 binding prote
27	102	5.9	1839	17	AAT18897	Human AL-1 cDNA.
28	101	5.8	636	16	AAQ85888	Human hek-L protei
29	73.2	4.2	3066	23	AAK57561	DNA encoding novel
30	61	3.5	374	22	AAI87907	Human polynucleoti
31	60.2	3.5	1631	22	AAH33111	Human colon cancer
32	59.4	3.4	1542	22	AAO08420	Human secreted pro
33	59.2	3.4	491	22	AAI87701	Human polynucleoti
34	59	3.4	836	22	AAI94700	Human neuroblastom
35	58.8	3.4	808	20	AAK89617	Human secreted pro
36	58.8	3.4	808	22	AAS59248	Human cDNA encodin
37	58.8	3.4	808	24	ABA90917	Human polynucleoti
38	58.8	3.4	1898	22	AAF30056	Human cDNA encodin
39	58.8	3.4	1899	22	AAC91465	Human PRO306 cDNA.
40	58.4	3.4	1564	20	AAK22231	Human secreted pro
41	58.2	3.3	418	22	AAI90717	Human polynucleoti
42	58.2	3.3	612	22	AAH71471	Human cervical can
43	58	3.3	641	21	AAK33356	Human secreted pro
44	58	3.3	694	22	AAO62000	Human uteroglobin-
45	57.8	3.3	735	22	AAO02202	Human reproductive

ALIGNMENTS

RESULT 1  
AAT03883  
ID AAT03883 standard; DNA; 1070 BP.  
AC AAT03883;  
DT 16-MAY-1996 (first entry)  
DE Eph transmembrane tyrosine kinase family ligand, Efl-2 encoding DNA.  
KW Efl-2; EHK1-L; Eph transmembrane tyrosine kinase family ligand;  
KW neurological disorder; identification; diagnosis; ss.  
OS Homo sapiens.  
PH Key Location/Qualifiers  
FT CDS 46..750  
FT /tag= a  
FT /product= Efl-2  
PN WO9527060-A2.  
XX 12-OCT-1995.  
XX 04-APR-1995; 95WO-US04208.  
XX 21-OCT-1994; 94US-0327423.  
XX 04-APR-1994; 94US-0222075.  
XX 12-APR-1994; 94US-0229402.  
XX 01-SEP-1994; 94US-0299567.  
XX (REGE-) REGENERON PHARM INC.

[illegible]

CC The sequence is that of a clone encoding hek-L protein, a protein  
CC that can bind hek (a cell surface receptor tyrosine kinase). Hek-L  
CC is the first known ligand for hek and can be used to study cellular  
CC processes regulated by hek (which may be involved in tumorigenesis)  
CC







	Best Local Similarity	63.3%	Pred. No. 9.4e-21;						
	Matches	276;	Conservative	0;	Mismatches	133;	Indels	27;	Gaps
Qy	16	gctgctgtgcccgtgcgcgtgtgcgcgtgtgcgcgaaggcccgagggcgctggg	75						
Db	109	gctgctgcgcgcatcgtgcgcgtctgctgtgagcgcagcccgcaagtgatctc	168						

169	ggaacgtacgcgtctcatggaaccgcagacacccacggttccaccgcgggattacac	228
Db		
136	cgtgcaggtaaactgaacactatctggattattactgccgcactacacagctcggg	195
Qy		
229	cgcgaggtgagcatcaatgactacctggacatctatgccctcaactacga-----	279
Db		
196	ggtgggccccgggcggggagccgggcccgcgagggcgagcgagcagtcgtgtgtacat	255
Qy		
280	-----ggagcgcgtgccgcgcgagcgcatggagcgctacgtctctctacat	324
Db		
256	ggtgagcgcgaacgcgtacctgcgcacctgcaacgcgcagcc-----aggcgttcagcgctggga	312
Qy		
325	ggtcaactacgagggccacgcgctctctggaccaccgcgagaaagggtctcaaacgcttggga	384
Db		

[illegible]

AAV42927	
ID	AAV42927 standard; cDNA; 783 BP.
XX	
AC	AAV42927;
XX	
XX	20-OCT-1998 (first entry)
DT	
XX	
XX	cDNA encoding an avian Elf-1 protein.
DE	
XX	
XX	Chicken; EPH receptor ligand; Elf-1; mek-4; sek-AP;
KW	tyrosine kinase ligand; B61; LERK-2; proliferation; differentiation;
KW	intracellular signalling; increased; survival; neuronal cell;
KW	neuron survival; treatment; Alzheimer's; parkinson's; lymphatic tumour;
KW	artificial liver; cartilage; bone formation; ss.
XX	
XX	

XX	Key	Location/Qualifiers
FH	5'UTR	1..85
FT		/*tag= a
FT		86..688
FT	CDS	/*tag= b

FT		/product= Elf-1
FT	sig_peptide	86..148
FT		/*tag= c
FT	3'UTR	686..783
FT		/*tag= d
XX		
PN	US5795734-A.	
XX		
PD	18-AUG-1998	

XX  
PF 31-MAY-1995: 95IIS-0455001 .

PR 31-MAY-1995; 95US-0455001.  
PR 19-SEP-1994; 94US-0308814.













GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2002, 08:42:05 : Search time 2227.85 Seconds  
(without alignments)  
16372.266 Million cell updates/sec

Title: US-09-733-756-1

Perfect score: 1743

Sequence: 1 gctgctgctgctgctgc.....aaaaaaaaaaaaaaaaaaaa 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	1678.8	96.3	1769	9	BC017722
2	1123.8	64.5	175826	2	AC021890
3	949.2	54.5	1070	6	AR004292
4	917.8	52.7	1037	6	AR001064
5	917.8	52.7	1037	6	AR080871
6	917.8	52.7	1037	6	AR164469
7	917.8	52.7	1037	6	AR120809
8	916	52.6	987	9	HSU14187
9	704	40.4	758	9	HUMERL2
c	643.2	36.9	289579	2	AC104632
11	642.2	36.8	326750	2	AC104327
12	253	14.5	1838	5	AB051678
13	241.8	13.9	355	10	AY045577
14	222	12.7	2535	10	U92885
c	183.4	10.5	186	6	AX32137
16	159.2	9.1	264	5	AF209777
17	149.2	8.6	642	6	AR134683
18	149.2	8.6	642	6	AR152044
19	149.2	8.6	665	5	CHKELF1LIG
20	149.2	8.6	783	6	AR023765
21	147.6	8.5	1066	9	HSA7292
22	139.6	8.0	1540	10	MMU14941
23	139.6	8.0	1594	10	MMU14752
24	139.6	8.0	1809	6	AR023764
25	135.4	7.8	314	6	AR134682
26	135.4	7.8	314	6	AR152043
27	135.4	7.8	4686	9	HS6LK02
28	135.4	7.8	36295	9	AC004258
29	130.2	7.5	799	5	GDRNARAGS
30	128.8	7.4	555	6	AR134677
31	128.8	7.4	555	6	AR152038
32	125.2	7.2	1650	5	DRTKRAL1
33	121.8	7.0	307	10	AF131912
c	121	6.9	162361	2	AC073818
35	119.8	6.9	651	5	XLU31204
36	119.4	6.9	452	5	XLU31205
37	119.4	6.9	2272	5	DRTKLELF1
38	119.2	6.8	188	10	MMU90666
39	118.2	6.8	981	5	AF317286
40	116.8	6.7	649	10	MUSMB61B
41	116.4	6.7	1480	6	AR048795
42	116.4	6.7	1480	6	AR103236
43	116.4	6.7	1480	6	I56902
44	116.4	6.7	1480	6	I88172
45	116.4	6.7	1480	9	HUMB61

## ALIGNMENTS

RESULT 1

BC017722

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC017722 Homo sapiens, ephrin-A3, clone MGC:21335 IMAGE:4397263, mRNA, 1769 bp linear PRI 06-DEC-2001 complete cds.

BC017722 Homo sapiens, ephrin-A3, clone MGC:21335 IMAGE:4397263, mRNA, 1769 bp linear PRI 06-DEC-2001 complete cds.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.

BC017722.1 GI:17389356 MGC. human.





```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="3"
/clone="RP11-498A2"
BASE COUNT      39638 a 46653 c 47059 g 40343 t 2133 others
ORIGIN

Query Match      64.5%; Score 1123.8; DB 2; Length 175826;
Best Local Similarity 99.8%; Pred. No. 7.7e-208;
Matches 1125; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 572 gaagactttgaggagagaaacctcaggtgcccgaagcttgaaagacatcagcgggacc 631
|||||
Db 39098 GAAGACTTTGAGGAGAGAAACCTCAGGTGCCCAAGCTTGAAAGAGCATCAGCGGGACC 39157

QY 632 agcccaacagggaacacgtgccctggcgtgggcatcgcttctctcatgacgttc 691
|||||
Db 39158 AGCCCCAAACGGGAACACCTGCCCTGGCGTGGGCATCGCCTTCTCTCATGACGTTT 39217

QY 692 ttggctctctagcttgcctccctcccttgggggagagatggcgggcttggaaag 751
|||||
Db 39218 TTGGCCTCTAGTCTGCCCCCTCCCTGGGGGGAGAGATGGGGCGGGCTTGGAAAG 39277

QY 752 agcaggagccttggcctctccaaagggaagcctagtgggctagagaccctctcccatg 811
|||||
Db 39278 AGCAGGAGCCCTTTGGCCTCTCCAAGGAAGCCTAGTGGGCTAGACCCCTCTCCCATG 39337

QY 812 gctagaagtgggctcgcaacatactgtgtccgcgccctctacccttcccccaag 871
|||||
Db 39338 GCTAGAAGTGGGCTGTCACCATACATCTGTGTCGCCCTCTACCCCTTCCGCCACG 39397

QY 872 tagggcactgtagtggacaagcagcgggagagccatgggtcccgggcggttgtggct 931
|||||
Db 39398 TAGGGCACTGTAGTGACCAAGCAGCGGGAGACGCACTGGTCCCGGGCGCTTTGGCT 39457

QY 932 ctgggtaatttggtaacaaacttggggcgaaggaaggaagtgactcaggactccctggc 991
|||||
Db 39458 CTGGTAATGTTGGTACCNAACTTGGGGGCCCAAAAGGGCAGTGTCTCAGGACTCCCTGGC 39517

QY 992 ccttggtaccttccctgactcctgtgtgcccttccctttgtcccccagagagacatat 1051
|||||
Db 39518 CCTTGTAACCTTTCCTGACTCTGCTGCTGCTCCTCTCCCTTTGTCCCTCCAGAGACATAT 39577

QY 1052 gccccagagagacaaatcaagcgtggaggaccccccaatgtctctctccaggggca 1111
|||||
Db 39578 GCCCCAGAGAGACAAATCAAGCGTGGGAGGACCCCCCAATGCTCTCTCCAGGGGCA 39637

QY 1112 gaacatggggagggaactagatgggcaaggggcagcactgctgtctcttctccctg 1171
|||||
Db 39638 GAACATGGGGAGGGACTAGATGGCAAGGGGACAGCACTGCCCTGCTCTTCCCTG 39697

QY 1172 ttacagcaataagcagctctctcccccactcccaactcccaagattgtggttggatt 1231
|||||
Db 39698 TTTACAGGCANTAAGCAGCTCTCTCTCCCTCCCACTCCCACTTCCAGGATGTGGTTGGATT 39757

QY 1232 gaaaccaagtttaacagttagacaccccttggggggcgggcagtggaacaaggatgccaaag 1291
|||||
Db 39758 GAAACCAAGTTTACAAGTAGACACCCCTGGGGGGCGGGCAGTGGACAGGATGGCAAG 39817

QY 1292 ggtggcaattgggtggcaggcagcaigtacagactctatatctctatatataatgtac 1351
|||||
Db 39818 GGTGGCAATTGGGTGGCCAGGCAGGCATGTACAGACTCTATATCTCTATATATAATGTAC 39877

QY 1352 agacagacagagtccttccctctttaaacccttgaccttcttgacttccctcagct 1411
|||||
Db 39878 AGACAGACAGAGTCCCTTCCCTTTTAACCCCTTGACCTTTCCTTGACTTTCCTTCAGCT 39937

QY 1412 tcagaccccttccccacagggttaggccccccacacactgggggacccccctggcccccttt 1471
|||||
Db 39938 TCAGACCCCTTCCCCACAGGCTAGGCCCCCCACACTGTGGGGACCCCTGGCCCTCTT 39997

QY 1472 ttgtcttctgtgaagacagagacctatgcaacgcacagacacatttttggagaccgttaaac 1531
|||||

```





QY	481	ctgtcgccctccacatcgcaactcgggggaaagcgggtcccacactctcccacagtccac	540
Db	577	CTGTGCGCCTCCACATCGCACATCCGGSGAGAGCGGTCCCCACATCTCCCCCAGTTTAC	636
QY	541	catggggcccaatgtgaagatcaacgtgctggaagactttgagggagagaacccctcaggt	600
Db	637	CATGGGCCCCAATGTGAAGATCAACGTGCTGGAAGACTTTGAGGGGAGAGAACCTCAGGT	696
QY	601	gcccaacttgagaagagcatcagcgggaccagcccccacacgggaacacactgcacctggc	660
Db	697	GCCCAAGCTTGAGAGAGCATCAGGGGACACGCCCCAAACGGGAGACACTGCCCCCTGGC	756
QY	661	ctggggcatcgcccttctctctcatgacgttcttggcctctagctctggccctccctcg	720
Db	757	CGTGGGCATCGCCTTCTTCCTCATGACGTTCTTGGGCTCCTAGCTCTGCCCTCCCGCTG	816
QY	721	ggggggagagatggggcggggcttgaaagagcaggagacctttggcctcttcaaggga	780
Db	817	GGGGGGAGAGATGGGGGGGGCTTGAAGAGACAGGAGCCCTTTGGCTCTCCAAGGGA	876
QY	781	agcctagtggccttagaccctctcccatggctagatgaagtggggcctgcaccatacatct	840
Db	877	AGCCTAGTGGGCTTAGACCCCTCTCCCATGGCTAGAAAGTGGGGCTCGACCATACATCT	936
QY	841	gtctcgccccctctacccttccccacgtaggcactgtagtgcacgaagcacggg	900
Db	937	GTGTCGCCCCCTCTACCCCTTCCCCCACGTAGGACACTGTAGTGGACCAAGCACGGGG	996
QY	901	acagccatgggtcccgggcg 921	
Db	997	ACAGCCATGGGTCCCGAGCAG 1017	

RESULT

8

HSU14187

LOCUS

DEFINITION

ACCESSION

U14187

VERSION

U14187.1

KEYWORDS

GI:642832

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Kozlosky,C.J., Maraskovsky,E., McGrew,J.T., VandenBos,T., Teepe,M., Lyman,S.D., Srinivasan,S., Fletcher,F.A., Gayle,R.B. III, Cerretti,D.P. and Beckmann,M.P.

TITLE

Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs encoding a family of proteins

JOURNAL

Oncogene 10 (2), 299-306 (1995)

REFERENCE

2

AUTHORS

Cerretti,D.P.

TITLE

Direct Submission

JOURNAL

Submitted (01-SEP-1994) Immunex Corp., 51 University St., Seattle, WA 98101, USA

FEATURES

source

1. .987

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1. .987

/gene="EPLG3"

58. .774

/gene="EPLG3"

/note="membrane-bound protein; glycosyl-phosphatidylinositol (GPI) anchored"

/codon\_start=1

/product="LTK-3"

/protein\_id="AAC50078.1"

/db\_xref="GI:642833"

/translation="MAAAPLLLLLVPLLPLLAQPGGALGNRHAVYWNSSQHL"

mat\_peptide

Best Local Similarity	100.0%;	Pred. NO. 9.8e-168;	
Matches	916;	Conservative	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;

Qy	1	gctgctgtcgtcgtcgtcgtcccgctgcgcgtcgtcgcgcgtcgtcgcgaaggcc	60
Db	72	GCTGCTGCTGCTGCTGCTGCTGCCGTGCCGTGCTGCCGTGCTGCSCCAAGGGC	131
Qy	61	cgaagggcactggaaaccggcatcgcggtgtactggaacagtccaaccacacactgcg	120
Db	132	CGGAGGGCCGCTGGAAACCGCATCGGTGTACTGGAAACAGTCCAACCAGCACCTGGG	191
Qy	121	gcgagagggctacaccgtgcagtgtaacgtgaacgactatctggattatttacctcccga	180
Db	192	GCAGAGSGCTACACCGTCAGGTGAACGTGAACGACTATCTTGATATTCTTGCCCGCA	251
Qy	181	ctacaacagctcgggggtggggccccggggcgagaccggggccgcgagcggggcagagca	240
Db	252	CTAACACAGCTCGGGGGTGGCCCCGGGGCGGACCGGGGCCGAGGGCGGACAGCA	311
Qy	241	gtacgtctgatcatgtagcgaacaggctaacgcacactgcacgcgcagccagggttl	300
Db	312	GTACGTGCTGTACATGGTAGCGCGCAACGGCTACCGCACTCTCAACCGCAGCACGGGCT	371
Qy	301	caagcgcgtggagtgcaaccgcgcgcacgcccgcacagccccatcaagttctcsgagaa	360
Db	372	CAAGCGCTGGAGTGCACCGCGCCACGCCCGCGACAGCCCCATCAAGTTCTCGGAGAA	431
Qy	361	gttcacagcgtacagcgctctctctgtgggttacgagttccacgcgcgcacagtagcta	420
Db	432	GTTCACAGCGTACAGCGCCTTCTCTGTGGCTACGAGTTCCACGCCGCGCACAGTACTA	491
Qy	421	ctacatctcacgcccaactcaaacctgcactggaagtgtctgagagatgaaggttgtcgt	480
Db	492	CTACATCTCCAGCCCCACTCAAACTGCACTTGAAGTGTCTGACGGATGAAGGTGTCGT	551
Qy	481	ctgctgcgctccacatcgactccgggagagaagccggtccccactctccccaggttcaa	540
Db	552	CTGCTCGCCTCCACATCGACTCCGGGGAGAAGCCGGTCCCCTCATCTCCCCAGTTTCA	611
Qy	541	catgggccccaatgtaagatcaacgctgctggaagactttgaggagagaagacctcaggt	600
Db	612	CATGGCCCCAATGTGAAGATCAAGTGAAGTGTGGAAGACTTTGAGGGAGAGAACCCTCAGST	671
Qy	601	qcccaagcttgagaagagacatcagcgggacacagccccaaaacgggaacacctgcacctggc	660
Db	672	GCCCAAGCTTGAGAAGAGCATCAGCGGGACCAAGCCCCAAAACGGGAACAACCTTGCCCTTGGC	731
Qy	661	cgtaggcatacgcttcttctctcatacgtctcttgccctccttagctctgcacctccctcctg	720
Db	732	CGTGGCATCGCCTTCTCCTCATCAGCTTCTTGCCCTCTTAGCTCTGCCCTTCCCCCTG	791
Qy	721	ggggggagagatggggcgggccttggaagagcaggagcctttggcctctccaaagga	780
Db	792	GGGGGGAGAGATGGGGCGGGCTTGGAAGAGCAGCAGGGAGCCTTTGGCCTCTCCAAGGGA	851
Qy	781	agccttagtgggctagacccctctcccatggctagagtggggcctgtcacatatcatct	840

```

QY 241 gtacgtgctgacatgtagcgcaacggtctacgcacactgcaacgcccagcagggtt 300
Db 261 GTACGTGCTGTACATGTTGAGCGCAACGGCTACCGACCTCAACGCCAGCGGGCTT 320
QY 301 caagcgtgagtagtgaacgcccgcacgcccgcacagcccatcaagttctcgagaa 360
Db 321 CAAGCGCTGGGAGTGAACGGCGCACGCCCGGCACAGCCCATCAAGTTCTCGGAA 380
QY 361 gtccagcgtacagcgcttctctggtctacgagttccacgcccagcagacta 420
Db 381 GTTCCAGCGCTACAGCGCTTCTCTGTGGCTTACAGTTCCAGCGCGCCAGGACTA 440
QY 421 ctacatctccagcccaactcaaacctgcaactggaagtgtctgagatgaaggtttcgt 480
Db 441 CTACATCTCCAGCGCCACTCAACCTGCACCTGGAAGTGTCTGAGGATGAAGTGTCTGT 500
QY 481 ctgctgagctccacatcgcaactcggggagagcgggtcccaactctccccagttcac 540
Db 501 CTGCTGCGCTCCACATCGCACTCGGGGAGAGCGGCTGCCACTCTCCCGCACTTCCAC 560
QY 541 catgggccccaaagtgaagatacaacgctgctgaagactttgaggagagacccctcaggt 600
Db 561 CATGGGCCCCAATGTGAGATCAAGCTGCTGGAACACTTTGAGGGAGAGAACCTTCCAGT 620
QY 601 gccaaagcttgagagagcagcagcgggagaccagcccccaacgggagacccctggc 660
Db 621 GCCAAGCTTGAGAGAGCATCAGCGGGAGACCGGCCCAACGGGAAACCTTCCCGCTGGC 680
QY 661 cgtggcctgctctctcctcctacagcgttctgctcctcctcctcctcctcctcctc 720
Db 681 CGTGGCCTGCGCTTCTCTCTCATACGCTTCTGCGCTCTCTGCTCTGCTCTGCTCTGCT 740
QY 721 gggggggagagatggggc 738
Db 741 GGGGGGAGAGATGGGC 758

RESULT 10
AC104632/c
LOCUS
DEFINITION AC104632 289579 bp DNA linear HTG 29-DEC-2001
SEQUENCE, 72 unordered pieces.
ACCESSION AC104632
VERSION AC104632.2 GI:17998594
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Hua, A. and Roe, B. A.
Mus musculus BAC Clone rp23-368d24
Unpublished
2 (bases 1 to 289579)
Hua, A. and Roe, B. A.
Direct Submission
Submitted (15-DEC-2001) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Dec 29, 2001 this sequence version replaced gi:17861055.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 72 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved.
* 1
* 2055: contig of 2055 bp in length
* 2155: gap of unknown length
* 2156: contig of 2080 bp in length
* 4235: gap of unknown length
* 4335: contig of 2191 bp in length
* 6526: gap of unknown length
* 6627: contig of 2828 bp in length
* 9454: gap of unknown length
* 9554: gap of unknown length
* 12042: contig of 2488 bp in length
* 12142: gap of unknown length
* 14176: contig of 2034 bp in length
* 14276: gap of unknown length
* 16599: contig of 2323 bp in length
* 16600: gap of unknown length
* 18844: contig of 2145 bp in length
* 18944: gap of unknown length
* 21106: contig of 2162 bp in length
* 21206: gap of unknown length
* 23193: contig of 1987 bp in length
* 23293: gap of unknown length
* 25652: contig of 2359 bp in length
* 25752: gap of unknown length
* 27729: contig of 1977 bp in length
* 27829: gap of unknown length
* 30718: contig of 2889 bp in length
* 30818: gap of unknown length
* 33823: contig of 3005 bp in length
* 33923: gap of unknown length
* 35961: contig of 2038 bp in length
* 36061: gap of unknown length
* 38732: contig of 2671 bp in length
* 38832: gap of unknown length
* 40937: contig of 2105 bp in length
* 41037: gap of unknown length
* 43594: contig of 2557 bp in length
* 43694: gap of unknown length
* 45995: contig of 2301 bp in length
* 46095: gap of unknown length
* 48083: contig of 1988 bp in length
* 48183: gap of unknown length
* 50630: contig of 2447 bp in length
* 50730: gap of unknown length
* 53982: contig of 3252 bp in length
* 54082: gap of unknown length
* 56167: contig of 2085 bp in length
* 56267: gap of unknown length
* 59028: contig of 2761 bp in length
* 59128: gap of unknown length
* 61945: contig of 2817 bp in length
* 62045: gap of unknown length
* 65116: contig of 3071 bp in length
* 65216: gap of unknown length
* 67626: contig of 2410 bp in length
* 67726: gap of unknown length
* 70795: contig of 3069 bp in length
* 70895: gap of unknown length
* 73334: contig of 2439 bp in length
* 73434: gap of unknown length
* 75894: contig of 2460 bp in length
* 75994: gap of unknown length
* 78387: contig of 2393 bp in length
* 78487: gap of unknown length
* 80841: contig of 2354 bp in length
* 80941: gap of unknown length
* 84112: contig of 3171 bp in length
* 84212: gap of unknown length
* 87107: contig of 2895 bp in length
* 87207: gap of unknown length
* 90387: contig of 3180 bp in length
* 90487: gap of unknown length
* 92854: contig of 2367 bp in length
* 92954: gap of unknown length

```

```

* 92955 96442: contig of 3488 bp in length
* 96443 96542: gap of unknown length
* 96543 98997: contig of 2455 bp in length
* 98998 99097: gap of unknown length
* 99098 102594: contig of 3496 bp in length
* 102594 102693: gap of unknown length
* 102693 105584: contig of 2891 bp in length
* 105584 105684: gap of unknown length
* 105684 108442: contig of 2758 bp in length
* 108442 108542: gap of unknown length
* 108542 111917: contig of 3375 bp in length
* 111917 112017: gap of unknown length
* 112017 116076: contig of 4059 bp in length
* 116076 116177: gap of unknown length
* 116177 119383: contig of 3207 bp in length
* 119383 119483: gap of unknown length
* 119483 123902: contig of 4419 bp in length
* 123902 124002: gap of unknown length
* 124002 127609: contig of 3607 bp in length
* 127609 127709: gap of unknown length
* 127709 131621: contig of 3912 bp in length
* 131621 131721: gap of unknown length
* 131721 135933: contig of 4212 bp in length
* 135933 136034: gap of unknown length
* 136034 140801: contig of 4768 bp in length
* 140801 140901: gap of unknown length
* 140901 144622: contig of 3721 bp in length
* 144622 147722: gap of unknown length
* 147722 148669: contig of 4147 bp in length
* 148669 148970: gap of unknown length
* 148970 153203: contig of 4234 bp in length
* 153203 153304: gap of unknown length
* 153304 157194: contig of 3891 bp in length
* 157194 157295: gap of unknown length
* 157295 162013: contig of 4719 bp in length
* 162013 162113: gap of unknown length
* 162113 167799: contig of 5686 bp in length
* 167799 167899: gap of unknown length
* 167899 173081: contig of 5182 bp in length
* 173081 173181: gap of unknown length
* 173181 180558: contig of 7377 bp in length
* 180558 180659: gap of unknown length
* 180659 186794: contig of 6136 bp in length
* 186794 186894: gap of unknown length
* 186894 192068: contig of 5174 bp in length
* 192068 192168: gap of unknown length
* 192168 198394: contig of 6226 bp in length
* 198394 198495: gap of unknown length
* 198495 203084: contig of 4590 bp in length
* 203084 203184: gap of unknown length
* 203184 208688: contig of 5504 bp in length
* 208688 208788: gap of unknown length
* 208788 214457: contig of 5669 bp in length
* 214457 214557: gap of unknown length
* 214557 219457: contig of 4899 bp in length
* 219457 219556: gap of unknown length
* 219556 226019: contig of 6463 bp in length
* 226019 226119: gap of unknown length
* 226119 234001: contig of 7882 bp in length
* 234001 234101: gap of unknown length
* 234101 241499: contig of 7398 bp in length
* 241499 241599: gap of unknown length
* 241599 249134: contig of 7535 bp in length
* 249134 249235: gap of unknown length
* 249235 259781: contig of 10547 bp in length
* 259781 259881: gap of unknown length
* 259881 267444: contig of 7563 bp in length
* 267444 267545: gap of unknown length
* 267545 279433: contig of 11888 bp in length
* 279433 279532: gap of unknown length
* 279532 289579: contig of 10047 bp in length
* 289579 289579: Location/Qualifiers
* 289579 1. 289579

```

FEATURES  
source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="rp23-388d24"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 70974 a 71815 c 70765 g 68626 t 7399 others
ORIGIN

Query Match 36.9%; Score 643.2; DB 2; Length 289579;
Best Local Similarity 81.3%; Pred. No. 1.4e-114;
Matches 920; Conservative 0; Mismatches 163; Indels 49; Gaps 13;

QY 572 gaagactttgaggagagaaacctcaggtgcccaagcttgagaagagcatcagcgagacc 631
|||||
Db 184805 GAAGACTTTGAGGAGAGAAATCCCGAGGTGCCAAGCTTGAGAGAGCATCATGCGGACC 184746

QY 632 agccccaacgggaaacacctgcccctggcctggcctggcctgcttctctctgaagttc 691
|||||
Db 184745 AGCCCCAAGCGGAACACCTGCTCTGGCGGTGGCATCGCTTCTTCATGACGCTC 184686

QY 692 ttggcctcctagctctgcccctcccctgggggggagagatgggggggcttggaaag 751
|||||
Db 184685 TTGGCCTCTCTAGCTCTGCCCTTCTCTATGACAGAGAGAGG--TGGAGAGGCTGAGAGG 184628

QY 752 agcagggagacctttggcctctccaaagggaagcctagtgggctagacacctctccatg 811
|||||
Db 184627 AGCAGGAGCTT-----GCTGTGGGGCTACATCCTTCTTCCCATG 184587

QY 812 gctagaagtggggcctgaccacacatctgtgtcgcgccccctct-----accctctcc 865
|||||
Db 184586 GTTGGGAGGGGGGTCTGCACCTGTACATCTCTCTGGGCTTGCCTCTTGGCCACACACTC 184527

QY 866 cccactgagggcactgtagtggaccaagcaggggagacgcatgggtcccggcgccctt 925
|||||
Db 184526 TCTAGCAGACGACCCGTAGTGGATCAGGCACAGGCACAGCCACAGGTCCCAGGTGCC-T 184468

QY 926 gtggctctggtaattttgggtaccaaacttggggggccaaaaggagcagtcaggaactc 985
|||||
Db 184467 GTGGCTTTGGTAATGTTGGTACCACAACTGGGGGCTATAA--GGCAGTGTCTAGGACTC 184409

QY 986 cctggccctggctaccttctcctgactcctggtgccccctctcctcttgcctcccccagag 1045
|||||
Db 184408 CTGGCCCCCTGGTACTCTTCCCTGACCTTGTGGCCCTCCCTCTTGTGCCCCAAGAGGC 184349

QY 1046 acatatgccccagagagagacaaatcgaaagcgtgggagaccccccatgtctctctcca 1105
|||||
Db 184348 AAATATG-CCCCAGAGAGAGAGAAA--CGAAGCATGGGAGGTGCCCTCTCCTCTCT-CT 184292

QY 1106 ggggcagaaatggggagggagactagatggcgaaggggcagcactgcctgcttctctt 1165
|||||
Db 184291 GGGCGCAGAACATGGGAGGGGACTAGGTGGGTGAGGGTGGAGCCTCAGGCTGC-CCCTC 184233

QY 1166 cccctgtttacagcaataagcagctctcctctccccactccactccactccaggttgggt 1225
|||||
Db 184232 CCCCTGTTTACAGCAATAAAGCATGCTCTCTCCCTCCACTCCACATAGAGAGACTGTGGTT 184173

QY 1226 tggattaaacccaagtttacaagtagacacccctggggggggcagcagtgacaaagatg 1285
|||||
Db 184172 TGGATTGAATCCAGTTTACAAATAGACACCCCTGGGGGGGAGCAGGAG-----TG 184121

QY 1286 ccaagggtgggcattgggtggcaggcagcagcatgtacagactctatatctctatatata 1345
|||||
Db 184120 GACAGGGGTGGGCATCGGGGTGCCAGACAGAGCATGTACAGACTCTATATCTCTATATA 184061

QY 1346 atgtacagacagacagagtcctctccctctttaaaccctcgaaccttcttgactccctc 1405
|||||
Db 184060 ATGTACAGACAGACAGAGTCCCTTCCCTCTTAAACCTTGACCTTCTCGACTTTCCTCTT 184001

QY 1406 tcagcttcagacctctcccccaccagcaggtgagccccccacacacacacacacacacacac 1465
|||||
Db 184000 CCAGACTCAGACCCCTTCCCTTACAGGTTAGGCCCC-----CCTTGGGACCCCTGGCC 183945

```

```
QY 1466 cctcttttctcttgaagacaggaactatgcaacgcacagacacacttttgagacagt 1525
|||||
Db 183944 CCTCTTTTGTCTCTGTGAGACAGAGACTATGCAACGCGACAGACACTTTTGGAGACCGT 183885
|||||
QY 1526 aaacaacacagcgcctccctccctccagccctgagcgcggaacatctccacaggaacttgc 1585
|||||
Db 183884 --AACAACAATGCCCTCCCTTCCAGCCCTGAGCGGGAACCAACTCCAGGACCTTGC 183827
|||||
QY 1586 cctgctcaccctatgtccacacatctcctcctcctggtcttttcaagtcttggctgt 1645
|||||
Db 183826 CTGCGCCACCTTAAGTGGTCC--CCACCCATCTCTGGGCTTTTCAAGTGTGGCTGT 183769
|||||
QY 1646 gactttcactctctcttctagctctaaaaataaactggagataaaaaa 1697
|||||
Db 183768 GACTTTCATACTCTCTTAGTCTAAAAATAAATAAAGTGAATAAAAAATA 183717
|||||

RESULT 11
AC104327
LOCUS
DEFINITION
AC104327
AC104327.2 GI:17978117
VERSION
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 326750)
Hua,A. and Roe,B.A.
Mus musculus BAC Clone rp23-295a4
Unpublished
2 (bases 1 to 326750)
Hua,A. and Roe,B.A.
Direct Submission
Submitted (10-DEC-2001) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Dec 24, 2001 this sequence version replaced gi:17439220.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 66 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1
2181: contig of 2180 bp in length
2281: gap of unknown length
2281: 4455: contig of 2175 bp in length
4456: gap of unknown length
4456: 7347: contig of 2792 bp in length
7348: gap of unknown length
7448: 10057: contig of 2610 bp in length
10058: 10157: gap of unknown length
10158: 12254: contig of 2097 bp in length
12255: 12354: gap of unknown length
12355: 14417: contig of 2063 bp in length
14418: 14517: gap of unknown length
14518: 18079: contig of 3562 bp in length
18080: 18179: gap of unknown length
18180: 20239: contig of 2060 bp in length
20240: 20339: gap of unknown length
20340: 22342: contig of 2003 bp in length
22343: 22442: gap of unknown length
22443: 24502: contig of 2060 bp in length
24503: 24602: gap of unknown length

24603:
24603: 25620: contig of 2018 bp in length
26621: gap of unknown length
26721: contig of 4056 bp in length
30776: gap of unknown length
30877: contig of 2312 bp in length
33188: gap of unknown length
33289: contig of 2728 bp in length
36016: gap of unknown length
36116: gap of unknown length
38176: contig of 2060 bp in length
38276: gap of unknown length
40782: contig of 2506 bp in length
40882: gap of unknown length
43672: contig of 2790 bp in length
43772: gap of unknown length
46580: contig of 2808 bp in length
46680: gap of unknown length
49679: contig of 2999 bp in length
49779: gap of unknown length
52528: contig of 2749 bp in length
52628: gap of unknown length
56785: contig of 4157 bp in length
56885: gap of unknown length
59118: contig of 2233 bp in length
59218: gap of unknown length
62964: contig of 3746 bp in length
63064: gap of unknown length
65910: contig of 2846 bp in length
66010: gap of unknown length
69396: contig of 3386 bp in length
69496: gap of unknown length
69497: contig of 3058 bp in length
72654: gap of unknown length
75720: contig of 3066 bp in length
75820: gap of unknown length
79364: contig of 3544 bp in length
79464: gap of unknown length
82599: contig of 3135 bp in length
82600: gap of unknown length
82699: contig of 2598 bp in length
85397: gap of unknown length
85398: contig of 3726 bp in length
89223: gap of unknown length
92441: contig of 3218 bp in length
92442: gap of unknown length
95856: contig of 3315 bp in length
95956: gap of unknown length
99508: contig of 3552 bp in length
99608: gap of unknown length
103789: contig of 4181 bp in length
103889: gap of unknown length
109457: contig of 5568 bp in length
109458: gap of unknown length
109557: gap of unknown length
114011: contig of 4454 bp in length
114111: gap of unknown length
114112: 118651: contig of 4540 bp in length
118652: gap of unknown length
118752: 123608: contig of 4857 bp in length
123609: gap of unknown length
123708: gap of unknown length
123709: 127713: contig of 4005 bp in length
127714: 127813: gap of unknown length
133303: contig of 5490 bp in length
133304: 133403: gap of unknown length
136767: contig of 3364 bp in length
136768: gap of unknown length
142660: contig of 5793 bp in length
142661: 142760: gap of unknown length
142761: 146701: contig of 3940 bp in length
146701: 146800: gap of unknown length
151040: contig of 4240 bp in length
151041: 151140: gap of unknown length
151141: 155319: contig of 4179 bp in length
155320: gap of unknown length
155447: contig of 4028 bp in length
```

```

* 159448 159547: gap of unknown length
* 159548 163961: contig of 4414 bp in length
* 163962 164061: gap of unknown length
* 164062 169232: contig of 5171 bp in length
* 169233 169332: gap of unknown length
* 169333 173172: contig of 3840 bp in length
* 173173 173272: gap of unknown length
* 173273 178980: contig of 5708 bp in length
* 178981 179080: gap of unknown length
* 179081 188181: contig of 9101 bp in length
* 188182 188281: gap of unknown length
* 188282 192848: contig of 4567 bp in length
* 192849 192948: gap of unknown length
* 192949 198349: contig of 5301 bp in length
* 198350 206031: contig of 7682 bp in length
* 206032 206131: gap of unknown length
* 206132 210282: contig of 4151 bp in length
* 210283 210382: gap of unknown length
* 210383 218415: contig of 8033 bp in length
* 218416 218515: gap of unknown length
* 218516 226120: contig of 7605 bp in length
* 226121 226220: gap of unknown length
* 226221 236580: contig of 10360 bp in length
* 236581 236680: gap of unknown length
* 236681 245273: contig of 8593 bp in length
* 245274 245373: gap of unknown length
* 245374 256115: contig of 10742 bp in length
* 256116 256215: gap of unknown length
* 256216 264078: contig of 7863 bp in length
* 264079 264178: gap of unknown length
* 264179 273637: contig of 9459 bp in length
* 273638 273737: gap of unknown length
* 273738 295230: contig of 21493 bp in length
* 295231 295330: gap of unknown length
* 295331 313298: contig of 17968 bp in length
* 313299 326750: contig of 13352 bp in length.
* 326750: Location/Qu:ifiers
* 326750: /organism="Mus musculus"
* 326750: /strain="C57BL/6J"
* 326750: /db_xref="taxon:10090"
* 326750: /clone="rp23-295a4"
* 326750: /clone_lib="RPCI mouse BAC library 23"
BASE COUNT 79344 a 80488 c 78974 g 81219 t 6725 others
ORIGIN

FEATURES
    source
        1..326750

Query Match      36.8%; Score 642.2; DB 2; Length 326750;
Best Local Similarity 81.2%; Pred. No. 2.3e-114;
Matches 919; Conservative 0; Mismatches 164; Indels 49; Gaps 13;

QY 572 gaagactttgaggagagaaacctcagtgcccaagcttgaaagacatcaggggacc 631
Db 195361 GAAGACTTTGAGGAGAGAACTCCAGTGCCCGCTGGCGCATCGCTTCTTCTCATGACGCTC 195420

QY 632 agccccaacgggaacacctgcccctggcgtgggcatcgcccttctctcatgacgttc 691
Db 195421 AGCCCCAAAGCGGAACACCTGCCCTGGCGCATCGCTTCTTCTCATGACGCTC 195480

QY 692 ttggcctctagctctgccccctccctgggggggagagatggggcggtctggaagg 751
Db 195481 TTGGCCTCTAGCTCTGCCCTTCTCTATGACAGAGAGAGG--TGGGAAGGCTGAGAAGG 195538

QY 752 agcaggagcctttggcctctccaagggaagccttagtgggcctagaccctctcccatg 811
Db 195539 AGCAGGGAGCTT-----GCTGTGGGCGCTACATCTTCTTCTTCCCATG 195579

QY 812 gctagaagtggcgctcacacatacatctgtgtccgccccctct-----acctctccc 865
Db 195580 GTTGGACGGGGGCTGCACGTGATACATCTCTCTGGGCGCTGCCCTCTTGTGCCACACACTC 195639

```

```

QY 866 cccacgtagggcactgtaggaccacgaacacacgggacacccatgggtcccgagcactt 925
Db 195640 TCTAGACAGGCACCCACTAGTGGATCAGGCACAGGACGCCACAGGTCCAGAGTGCC-T 195698

QY 926 gtggtctgtgtaattgtttgtaccacaaacttggggggcccaaaagggcagtgctcagactc 985
Db 195699 GTGGCTTTGGTAATGTTTGGTACCAAAACCTGGGGGCTATAAA--GGCAGTGCTCAGGACTC 195757

QY 986 cctggccccctgtacctttccctgactcctgtgtgccccctcctcctttgtccccccagagag 1045
Db 195758 CTTGGCCCCCTGGTACCTTTTCCCTTGACCTTTGGTGCCCTCCCTCTTTGTCCCCCAAGAGCG 195817

QY 1046 acatatgccccccagagagacaaatcgaagcgtgggagagcaccccccaattgctctctcca 1105
Db 195818 AAATATG-CCCCAGAGAGAAACAAA-CGAAGCATGGGAGGTGCCCTCTCTCTCTCT-CT 195874

QY 1106 ggggcagaacatggggaggggactagatgggcaaggggcagcactgctgtctcctt 1165
Db 195875 GGGGCAGAAACATGGGGAGGGGACTAGTGGGTGGAGGCTGAGGCTCAGGNTGC-CCCTC 195933

QY 1166 cccctgtttacagcaataagcagctcctcctccctccctccacactcccaactccagattgtgt 1225
Db 195934 CCCCTGTTTACAGCAATAAGCATGTCTCTCCCTCCACTCCACATAGAGGACTGTGGTT 195993

QY 1226 tggattgaacccaagtattacagtagacacccctggggggcgggcagtggaacaaggatg 1285
Db 195994 TGGATTGAATCCAAAGTTTACAAATAGACACCCCTGGGGGGAGCAGGAG-----TG 196045

QY 1286 ccaagggtggcattgggggtggcagcagcagcagcagcagcagcagcagcagcagcagcag 1345
Db 196046 GACAGGGGTGGGCATCGGGGTGCCAGACAGGAGCATGTACAGACTCTATATATCTCTATATATA 196105

QY 1346 atgtacagacagacagagatccctctcctctctttaaaccctcgacctttcttgacttccccct 1405
Db 196106 ATGTACAGACAGACAGACTCCCTTCCCTCTTCTTAACCTTTGACCTTTCTCGACTTCCCTCT 196165

QY 1406 tcaacttcagacccttccccacacaggtagggccccccacacacctgggggagccccctggcc 1465
Db 196166 CCAGACTCAGACCCCTTCCCTTACAGGTAGGGCCCC-----CCTTGGGACCCCTTGCCC 196221

QY 1466 cctctttgtctctgtgaagacagggacctatgcaacgacacagacacacttttgggagaccgt 1525
Db 196222 CCTCTTTTGTCTTCTGTGAGACAGAGGACTATGCAACGACACAGACACTTTTGGAGACCGT 196281

QY 1526 aaacaacagcgccccctcctcctcagccttgagcccggggaaacctctccagggacctgac 1585
Db 196282 --AACACAAATGCCCCCTCCTTCCAGCCCTGAGCGGGAACCAACTCCCAAGGACCTTGC 196339

QY 1586 cctgtcacctatgtgtggtccacctatcctcctgggccccttttcaagtgtgttggctgt 1645
Db 196340 CTTGCCACCCCTAAGTGGTCC--CCACCCATCCTGGGCTTTTCAAGTGTCTTGGCTGT 196397

QY 1646 gactttacactctgctcttagtctaaataaaataaaactggagataaaaaata 1697
Db 196398 GACTTTACACTCTGCTCTTAGTCTAAATAAAATAAACTGGAGATAAAAAATA 196449

```

## RESULT 12

AB051678

LOCUS

DEFINITION

AB051678

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AB051678 1838 bp mRNA linear VRT 20-SEP-2001  
 Danio rerio mRNA for ephrin-A3, complete cds.

AB051678.1 GI:14196230  
 Danio rerio cDNA to mRNA.

Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 1838)  
 Hirata, Y., Mieda, M., Harada, T., Yamasu, K. and Okamoto, H.  
 Identification of ephrin-A3 and novel genes specific to the  
 midbrain-MHB in embryonic zebrafish by ordered differential display





Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	949.2	54.5	1070	1	US-08-239-567-8	Sequence 8, Appli
2	917.8	52.7	1033	1	US-08-239-567-8	Sequence 1, Appli
3	917.8	52.7	1037	1	US-08-239-567-8	Sequence 1, Appli
4	917.8	52.7	1037	2	US-09-057-121-1	Sequence 1, Appli
5	917.8	52.7	1037	2	US-09-058-734-1	Sequence 1, Appli
6	149.2	8.6	642	1	US-09-609-324A-9	Sequence 9, Appli
7	149.2	8.6	642	2	US-08-920-440B-9	Sequence 9, Appli
8	149.2	8.6	642	4	US-09-173-492-9	Sequence 9, Appli
9	149.2	8.6	642	4	US-09-173-133-9	Sequence 9, Appli
10	149.2	8.6	642	1	US-08-455-001-3	Sequence 3, Appli
11	149.2	8.6	783	1	US-08-308-814-1	Sequence 1, Appli
12	139.6	8.0	1615	4	PCT-US95-11869-3	Sequence 1, Appli
13	139.6	8.0	1809	1	US-08-455-001-1	Sequence 1, Appli
14	139.6	8.0	1809	5	PCT-US95-11869-1	Sequence 1, Appli
15	135.4	7.8	314	1	US-09-609-324A-7	Sequence 7, Appli
16	135.4	7.8	314	4	US-08-920-440B-7	Sequence 7, Appli
17	135.4	7.8	314	4	US-09-173-492-7	Sequence 7, Appli
18	135.4	7.8	314	4	US-09-173-133-7	Sequence 7, Appli
19	135.4	7.8	314	4	US-09-165-533-7	Sequence 7, Appli
20	128.8	7.4	555	1	US-09-609-324A-1	Sequence 7, Appli
21	128.8	7.4	555	2	US-08-920-440B-1	Sequence 7, Appli
22	128.8	7.4	555	4	US-09-173-492-1	Sequence 7, Appli
23	128.8	7.4	555	4	US-09-173-133-1	Sequence 7, Appli
24	128.8	7.4	555	4	US-09-165-533-1	Sequence 7, Appli
25	128.8	7.4	555	5	PCT-US95-12779-1	Sequence 7, Appli
26	128.8	7.4	555	5	PCT-US95-15781-1	Sequence 7, Appli
27	116.4	6.7	1480	1	US-08-448-736-11	Sequence 11, Appli

Query Match	54.5%;	Score 949.2;	DB 1;	Length 1070;
Best Local Similarity	97.0%;	Pred. No. 9e-186;		
Matches 976;	Conservative 0;	Mismatches 18;	Indels 12;	Gaps 1;

  

yy	1	gcctgctgtcgtcgtcgtcgctgcccgctccgctgctgcgcgtgctgcggccaaggccc	60
yb	60	GCTGCTGCTGCTGCTGCTGCTGCCCGTGCCTGCTGCTGCTGCTGCCCCAAGGGCC	119













```

Db      487  TGCCTGCGACTGAAGGTGTACGTGCGCGCCGACCAACGAGACCCCTGTGTACGAGG 538
      II IIII I IIIIIIIII IIII I IIII I IIII I IIII I IIII I IIII I
RESULT 10
US-08-455-001-3
; Sequence 3, Application US/08455001
; Patent No. 5795734
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Therefo
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,001
; FILING DATE: 31 May 1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709

```

/	TELEPHONE:	(617) 227-7400
/	TELEFAX:	(617) 227-5941
/	INFORMATION FOR SEQ ID NO:	3:
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	783 base pairs
/	TYPE:	nucleic acid
/	STRANDEDNESS:	both
/	TOPOLOGY:	linear
/	MOLECULE TYPE:	cdna
/	FEATURE:	
/	NAME/KEY:	CDS
/	LOCATION:	86..685
/	FEATURE:	
/	NAME/KEY:	5'UTR
/	LOCATION:	1..85
/	FEATURE:	
/	NAME/KEY:	3'UTR
/	LOCATION:	686..783
/	FEATURE:	
/	NAME/KEY:	sig peptide
/	LOCATION:	86..148
/	US-08-455-001-3	
Query Match                  8.6%; Score 149.2; DB 1; Length 783;		
Best Local Similarity      63.3%; Pred. No. 4.2e-22;		
Matches 276; Conservative         0; Mismatches 133; Indels    27; Gaps		
/QY	16	gctgctgcgcccgtagccgctgctgcccgaaggccggagggcgctggg 75 
/Ddb	109	GCTGCTCGCCGCATGTGGCGTCTGCGGTGTGGAGCAGCACCCCAGCAAGTGATCTC 168
/QY	76	aaccggcgtcgcgtgtaactggaaacagctccaaccagcacctgcggcgagagggtacac 135 
/Ddb	169	GGACCGGTAGCGCGCTCTATTTTGAACCGCAGCAACCACAGGTTCCACCGCGGGATTACAC 238
/QY	136	cgtcaggtgaaagtgaacgacctatctgatatttactgccgcgcactacaacagctcgg 195 





Query Match 7.88; Score 135.4; DB 1; Length 314;

123	Qy	gagagggctacacccgtgcaggtgaacgtgaacgactatctggatatctactgccgcgact	181
124	Qy		
125	Qy		
126	Qy		
127	Db	gcgggggctacacgttggagctgagatcaatgactacctgacatctacttgcgcgact	86
128	Db		
129	Qy		
130	Qy		
131	Qy		
132	Qy		
133	Qy		
134	Qy		
135	Qy		
136	Qy		
137	Qy		
138	Qy		
139	Qy		
140	Qy		
141	Qy		
142	Qy		
143	Qy		
144	Qy		
145	Qy		
146	Qy		
147	Qy		
148	Qy		
149	Qy		
150	Qy		
151	Qy		
152	Qy		
153	Qy		
154	Qy		
155	Qy		
156	Qy		
157	Qy		
158	Qy		
159	Qy		
160	Qy		
161	Qy		
162	Qy		
163	Qy		
164	Qy		
165	Qy		
166	Qy		
167	Qy		
168	Qy		
169	Qy		
170	Qy		
171	Qy		
172	Qy		
173	Qy		
174	Qy		
175	Qy		
176	Qy		
177	Qy		
178	Qy		
179	Qy		
180	Qy		
181	Qy		
182	Qy		
183	Qy		
184	Qy		
185	Qy		
186	Qy		
187	Qy		
188	Qy		
189	Qy		
190	Qy		
191	Qy		
192	Qy		
193	Qy		
194	Qy		
195	Qy		
196	Qy		
197	Qy		
198	Qy		
199	Qy		
200	Qy		
201	Qy		
202	Qy		
203	Qy		
204	Qy		
205	Qy		
206	Qy		
207	Qy		
208	Qy		
209	Qy		
210	Qy		
211	Qy		
212	Qy		
213	Qy		
214	Qy		
215	Qy		
216	Qy		
217	Qy		
218	Qy		
219	Qy		
220	Qy		
221	Qy		
222	Qy		
223	Qy		
224	Qy		
225	Qy		
226	Qy		
227	Qy		
228	Qy		
229	Qy		
230	Qy		
231	Qy		
232	Qy		
233	Qy		
234	Qy		
235	Qy		
236	Qy		
237	Qy		
238	Qy		
239	Qy		
240	Qy		
241	Qy		
242	Qy		
243	Qy		
244	Qy		
245	Qy		
246	Qy		
247	Qy		
248	Qy		

Search completed: July 13, 2002, 09:49:40  
Job time: 4015 sec